

Oklahoma School Testing Program
Oklahoma Core Curriculum Tests
Technical Report
Spring 2010 Administration

August, 2010

Table of Contents

INTRODUCTION.....	1
CHAPTER I. OVERVIEW OF THE OCCT	2
1.1 Skills Assessed by the OCCT	2
1.2 Test Development Procedures	3
1.3 Configuration of the Tests	3
CHAPTER II. STATISTICAL ANALYSIS.....	23
2.1 Data Files for Statistical Analysis.....	23
2.2 Analysis of the Multiple-Choice Tests	23
2.2.1 Classical Item Analyses	23
2.2.2 Differential Item Functioning Analyses.....	26
2.2.3 Item Calibration and Equating	28
2.2.4 Raw Score to Scaled Score Conversion.....	30
2.2.5 Test Score Reliability.....	30
Coefficient Alpha.....	30
Standard Error of Measurement.....	31
Conditional Standard Error of Measurement.....	32
Reliability of Performance-Level Classification Decisions.....	32
2.2.6 Validity	34
2.3 Analysis of the Writing Tests	40
2.3.1 Prompt Scoring Formula.....	40
2.3.2 Statistical Adjustments to Scale the Writing Scores.....	41
Adjustment for Prompt Difficulty and Rater-Year Effects.....	41
Adjustment for Rater-Year Effects	41
A Compound Adjustment	42
2.3.3 Rater Agreement for Operational Writing Prompts.....	44
CHAPTER III. STATE RESULTS.....	46
CHAPTER IV. PERFORMANCE STANDARDS	67
REFERENCES.....	68
APPENDIX A. DATA REVIEW RESULTS.....	70
APPENDIX B. RAW-TO-SCALED SCORE CONVERSION TABLES AND FREQUENCY DISTRIBUTIONS	72

List of Tables

Table 1.3.A	
Number of Operational and Field-test Items by Content Area and Grade	3
Table 1.3.B	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 3 Reading	4
Table 1.3.C	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 4 Reading	5
Table 1.3.D	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 5 Reading	6
Table 1.3.E	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 6 Reading	7
Table 1.3.F	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 7 Reading	8
Table 1.3.G	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 8 Reading	9
Table 1.3.H	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 3 Mathematics	10
Table 1.3.I	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 4 Mathematics	11
Table 1.3.J	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 5 Mathematics	12
Table 1.3.K	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 6 Mathematics	13
Table 1.3.L	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 7 Mathematics	14
Table 1.3.M	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 8 Mathematics	15
Table 1.3.N.1	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 5 Science Process Standards	16
Table 1.3.N.2	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 5 Science Content Standards	17
Table 1.3.O.1	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 8 Science Process Standards	18

Table 1.3.O.2	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 8 Science Content Standards.....	19
Table 1.3.P	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 5 Social Studies.....	20
Table 1.3.Q	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 7 Geography.....	21
Table 1.3.R	
2010 <i>PASS</i> Blueprint and Actual Item Counts: Grade 8 U.S. History	22
Table 2.2.A	
Classical Item and Test Analyses Summary for the Operational Forms	25
Table 2.2.B	
Item Analyses Summary for Field-test Items	26
Table 2.2.C	
Number of Anchor Items by Subject and Grade.....	30
Table 2.2.D	
Conditional Standard Errors of Measurement for Each Achievement Level Cut Score	33
Table 2.2.E	
Estimates of the Reliability of Decisions for Specified Cut Scores ^a	34
Table 2.2.F	
Standards Intercorrelation: Grade 3 Reading.....	36
Table 2.2.G	
Standards Intercorrelation: Grade 4 Reading.....	36
Table 2.2.H	
Standards Intercorrelation: Grade 5 Reading.....	36
Table 2.2.I	
Standards Intercorrelation: Grade 6 Reading.....	36
Table 2.2.J	
Standards Intercorrelation: Grade 7 Reading.....	36
Table 2.2.K	
Standards Intercorrelation: Grade 8 Reading.....	37
Table 2.2.L	
Standards Intercorrelation: Grade 3 Mathematics	37
Table 2.2.M	
Standards Intercorrelation: Grade 4 Mathematics	37

Table 2.2.N	
Standards Interrelation: Grade 5 Mathematics	38
Table 2.2.O	
Standards Interrelation: Grade 6 Mathematics	38
Table 2.2.P	
Standards Interrelation: Grade 7 Mathematics	38
Table 2.2.Q	
Standards Interrelation: Grade 8 Mathematics	38
Table 2.2.R	
Standards Interrelation: Grade 5 Science.....	39
Table 2.2.S	
Standards Interrelation: Grade 8 Science.....	39
Table 2.2.T	
Standards Interrelation: Grade 5 Social Studies	39
Table 2.2.U	
Standards Interrelation: Grade 7 Geography	39
Table 2.2.V	
Standards Interrelation: Grade 8 U. S. History	40
Table 2.3.A	
Weights Assigned to Writing Analytic Traits.....	41
Table 2.3.B	
Scaled Score Ranges for Each Achievement Level	42
Table 2.3.C	
Sample Means and Standard Deviations Used for Calculating Constants	43
Table 2.3.D	
Grade 5 Writing Results.....	43
Table 2.3.E	
Grade 8 Writing Results.....	44
Table 2.3.F	
Inter-rater Agreement for Operational Writing Prompts	45
Table 3.1	
Means and Standard Deviations of Students' Raw Scores and Scaled Scores.....	46
Table 3.2	
Percentage of Students Performing within Each Achievement Category in 2008 to 2010	47

Table 3.3	
Subgroup Results: Grade 3 Reading	48
Table 3.4	
Subgroup Results: Grade 4 Reading	49
Table 3.5	
Subgroup Results: Grade 5 Reading	50
Table 3.6	
Subgroup Results: Grade 6 Reading	51
Table 3.7	
Subgroup Results: Grade 7 Reading	52
Table 3.8	
Subgroup Results: Grade 8 Reading	53
Table 3.9	
Subgroup Results: Grade 3 Mathematics.....	54
Table 3.10	
Subgroup Results: Grade 4 Mathematics.....	55
Table 3.11	
Subgroup Results: Grade 5 Mathematics.....	56
Table 3.12	
Subgroup Results: Grade 6 Mathematics.....	57
Table 3.13	
Subgroup Results: Grade 7 Mathematics.....	58
Table 3.14	
Subgroup Results: Grade 8 Mathematics.....	59
Table 3.15	
Subgroup Results: Grade 5 Science	60
Table 3.16	
Subgroup Results: Grade 8 Science	61
Table 3.17	
Subgroup Results: Grade 5 Social Studies.....	62
Table 3.18	
Subgroup Results: Grade 7 Geography.....	63
Table 3.19	
Subgroup Results: Grade 8 History	64

Table 3.20	
Subgroup Results: Grade 5 Writing.....	65
Table 3.21	
Subgroup Results: Grade 8 Writing.....	66
Table 4.1 Final Scaled Score Ranges for Reading and Mathematics	67
Table A	
Data Review Results	71
Table B.1	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 3 Reading	73
Table B.2	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 4 Reading	74
Table B.3	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Reading	75
Table B.4	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 6 Reading	76
Table B.5	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 7 Reading	77
Table B.6	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Reading	78
Table B.7	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 3 Mathematics	79
Table B.8	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 4 Mathematics	80
Table B.9	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Mathematics	81
Table B.10	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 6 Mathematics	82
Table B.11	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 7 Mathematics	83
Table B.12	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Mathematics	84
Table B.13	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Science	85
Table B.14	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Science	86

Table B.15	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Social Studies	87
Table B.16	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 7 Geography	89
Table B.17	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 U.S. History.....	90
Table B.18	
Composite Score Frequency Distribution: Grade 5 Writing.....	91
Table B.19	
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Writing	92

INTRODUCTION

The Oklahoma Core Curriculum Test (OCCT) is a component of the Oklahoma School Testing Program (OSTP) administered in Grades 3 through 8. It is a transparent, standard-based, criterion-referenced assessment system designed to monitor student achievement of the Oklahoma *Priority Academic Student Skills (PASS)* adopted by the Oklahoma State Board of Education. Currently, the OCCT includes direct Writing assessments in Grades 5 and 8 and Multiple-Choice (MC) assessments of Reading and Mathematics in Grades 3 through 8; Science in Grades 5 and 8; as well as Social Studies in Grade 5; Geography in Grade 7; and U.S. History, Constitution, and Government in Grade 8.

In 2010, the OCCT was administered during the spring. The Writing assessments were administered on February 24th and March 3rd. The Grade 7 Geography and Grade 8 Reading and Mathematics were administered online during an April 12th to May 14th window. The remaining tests were administered during the MC testing window of April 12th – May 7th.

This technical report outlines the statistical analyses that were carried out in support of the 2010 OCCT. Chapter I provides an overview of the test content and design. Chapter II details the statistical procedures that were carried out in support of the OCCT. These procedures include preliminary item analyses, differential item functioning analyses, calibration and equating, and various miscellaneous analyses. Chapter III presents statewide test results. Chapter IV describes the performance standard setting process and results. Two appendices are provided. Appendix A presents the data review results. Appendix B presents the raw score to scaled score (RS – SS) conversion tables and frequency distribution by grade.

The technical information provided in this report is intended for use by all interested in how the test is evaluated, how the scores are interpreted, and the subsequent educational decisions based on the test results. It is assumed that the reader has technical knowledge of test construction and measurement procedures, as stated in *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999).

CHAPTER I. OVERVIEW OF THE OCCT

The purpose of the Oklahoma Core Curriculum Test (OCCT) is to fulfill accountability requirements and to provide feedback about student mastery of the knowledge and skills delineated in the Oklahoma *Priority Academic Student Skills (PASS)* standards. In the spring of 2010, the OCCT assessments were administered to all eligible public school students in Grades 3 through 8. The OCCT includes assessments of Reading and Mathematics in Grades 3 through 8; Writing and Science in Grades 5 and 8; as well as Social Studies in Grade 5; Geography in Grade 7; and U.S. History, Constitution, and Government in Grade 8. All tests were designed to measure the Oklahoma *Priority Academic Student Skills (PASS)* adopted by the Oklahoma State Board of Education. The 2010 administration of the OCCT was the sixteenth for students in Grades 5 and 8 and the sixth for students in Grades 3, 4, and 7 (Geography only). This was the fifth operational administration of the Reading and Mathematics tests in Grades 6 and 7.

Data Recognition Corporation (DRC) worked with the Oklahoma State Department of Education (SDE) to construct OCCT test forms aligned to the *PASS* standards. Each test form included a set of operational items used to produce student test scores, and a set of embedded field-test items. The Writing assessments included one extended constructed-response (CR) item. The Reading, Mathematics, Science, Social Studies, Geography and History assessments were composed of Multiple-Choice (MC) items only. For each content and grade, there were eight forms consisting of a common set of operational items and a unique set of 10 field-test items. Responses to the operational items were used to produce student scores. Responses to the field-test items were used to evaluate the psychometric properties of these newly developed items for possible inclusion on future forms.

The OCCT is an untimed test. The MC tests in Grades 3 – 5 were administered in two sessions. The Writing tests in Grades 5 and 8 and the MC tests in Grades 6 – 8 were administered in one session. With the exception of Grade 7 Geography and Grade 8 Reading and Mathematics, all assessments were administered as paper-and-pencil tests. The Grade 7 Geography and Grade 8 Reading and Mathematics assessments were delivered primarily online, with paper forms available for accommodated administrations and for make-ups.

In the following sections, more information is provided on the skills assessed by the OCCT, test development procedures, and the configuration of the tests.

1.1 Skills Assessed by the OCCT

The standards assessed by each test can be found at <http://sde.state.ok.us/Curriculum/PASS/default.html>. The OCCT measures all *PASS* standards except for objectives that cannot be appropriately measured within the limitations of a large scale, Multiple-Choice test. For example, the majority of *PASS* standards (listening, reviewing, etc.) are not measured in the ELA assessment. Standards not measured by the OCCT must be assessed by local school districts.

1.2 Test Development Procedures

The items used in the operational 2010 OCCT were selected from the SDE-owned pool of items. All items selected had previously been reviewed and approved by Oklahoma content, bias, and sensitivity review committees. These operational items had been field tested during previous administrations. The field-test statistics for these operational items indicated that the items were of acceptable quality.

For field-test items embedded in the OCCT, DRC assessment specialists selected field test ready items from SDE's item bank, as well as items newly-developed for 2010. A total of 80 items per content/grade were selected for embedded field-testing.

1.3 Configuration of the Tests

Table 1.3.A shows the number of operational and field-test items by content area and grade used in the 2010 operational tests. Also shown is the number of operational items included in the anchor set used for equating the 2010 forms to the previously established reporting scale. For all Multiple-Choice tests, each form contained a core set of operational items common across forms and a unique set of field-test items.

Table 1.3.A
Number of Operational and Field-test Items by Content Area and Grade

Content Area	Grade	Number of Forms	Number of Operational Items per Form ^a	Number of Operational Items in Anchor Set ^a	Number of Field-test Items per Form	Total Number of Items per Form	Total Number of Field-test Items Per Grade
Reading	3-8	8	50	14-18 ^b	10	60	80
Writing	5, 8	1	1	0	0	1	0
Mathematics	3-8	8	45	17-20 ^c	10	55	80
Science	5, 8	8	45	16,14	10	55	80
Social Studies	5	8	60	20	10	70	80
Geography	7	8	45	15	10	55	80
U.S. History	8	8	45	15	10	55	80

^a Operational item counts include anchor items.

^b Anchor counts for Reading tests were 17, 18, 16, 14, 17, and 15 in Grades 3 through 8 respectively.

^c Anchor counts for Mathematics tests were 20, 19, 18, 18, 17, and 17 in Grades 3 through 8 respectively.

Tables 1.3.B through 1.3.R provide information drawn from the official 2010 test *PASS* blueprints. These tables show the number of items by content standard specified in the blueprints and the number of items that appeared on the 2010 operational assessments.

Table 1.3.B
2010 *PASS* Blueprint and Actual Item Counts: Grade 3 Reading

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i>*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Vocabulary	12	12	24%
Words in Context (2.1)	2-4	3	
Affixes, Roots, and Stems (2.2)	2-4	3	
Synonyms, Antonyms, and Homonyms (2.3)	2-4	2	
Using Resource Materials (2.4)	2-4	4	
Comprehension/Critical Literacy	24	24	48%
Literal Understanding (4.1)	5	5	
Inferences and Interpretation (4.2)	7	7	
Summary and Generalization (4.3)	6	6	
Analysis and Evaluation (4.4)	6	6	
Literature	8	8	16%
Literary Elements (5.2)	4	2	
Figurative Language/Sound Devices (5.3)	4	6	
Research and Information	6	6	12%
Accessing Information (6.1)	6	6	
<i>Total Test</i>	50	50	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.C
2010 PASS Blueprint and Actual Item Counts: Grade 4 Reading

PASS Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Vocabulary	12	12	24%
Words in Context (1.1)	4	3	
Affixes, Roots, and Stems (1.2)	4	5	
Synonyms, Antonyms, and Homonyms (1.3)	4	4	
Comprehension/Critical Literacy	23	23	46%
Literal Understanding (3.1)	4	4	
Inferences and Interpretation (3.2)	6	6	
Summary and Generalization (3.3)	7	7	
Analysis and Evaluation (3.4)	6	6	
Literature	9	9	18%
Literary Elements (4.2)	5	5	
Figurative Language/Sound Devices (4.3)	4	4	
Research and Information	6	6	12%
Accessing Information (5.1)	6	6	
Total Test	50	50	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.D
2010 PASS Blueprint and Actual Item Counts: Grade 5 Reading

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Vocabulary	12	12	24%
Words in Context (1.1)	4	4	
Affixes, Roots, and Stems (1.2)	4	4	
Synonyms, Antonyms, and Homonyms (1.3)	4	4	
Comprehension/Critical Literacy	20	20	40%
Literal Understanding (3.1)	4	4	
Inferences and Interpretation (3.2)	4-6	6	
Summary and Generalization (3.3)	4-6	5	
Analysis and Evaluation (3.4)	4-6	5	
Literature	12	12	24%
Literary Genre (4.1)	4	4	
Literary Elements (4.2)	4	4	
Figurative Language/Sound Devices (4.3)	4	4	
Research and Information	6	6	12%
Accessing Information (5.1)	2-4	4	
Interpreting Information (5.2)	2-4	2	
Total Test	50	50	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.E
2010 PASS Blueprint and Actual Item Counts: Grade 6 Reading

PASS Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Vocabulary	8	8	16%
Words in Context (1.1)	4	4	
Word Origins (1.2)	4	4	
Comprehension/Critical Literacy	20	20	40%
Literal Understanding (3.1)	4	4	
Inferences and Interpretation (3.2)	4-6	6	
Summary and Generalization (3.3)	4-6	6	
Analysis and Evaluation (3.4)	4-6	4	
Literature	14	14	28%
Literary Genres (4.1)	4	4	
Literary Elements (4.2)	4-6	5	
Figurative Language/Sound Devices (4.3)	4-6	5	
Research and Information	8	8	16%
Accessing Information (5.1)	4	4	
Interpreting Information (5.2)	4	4	
Total Test	50	50	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.F
2010 PASS Blueprint and Actual Item Counts: Grade 7 Reading

PASS Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Vocabulary	10	10	20%
Words in Context (1.1)	3-4	4	
Word Origins (1.2)	3-4	3	
Idioms and Comparisons (1.3)	3-4	3	
Comprehension/Critical Literacy	20	20	40%
Literal Understanding (3.1)	4	4	
Inferences and Interpretation (3.2)	4-6	6	
Summary and Generalization (3.3)	4-6	6	
Analysis and Evaluation (3.4)	4-6	4	
Literature	12	12	24%
Literary Genres (4.1)	4	4	
Literary Elements (4.2)	4	4	
Figurative Language/Sound Devices (4.3)	4	4	
Research and Information	8	8	16%
Accessing Information (5.1)	4	4	
Interpreting Information (5.2)	4	4	
Total Test	50	50	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.G
2010 PASS Blueprint and Actual Item Counts: Grade 8 Reading

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Vocabulary	6	6	12%
Words in Context (1.1)	2	3	
Word Origins (1.2)	2	1	
Idioms and Comparisons (1.3)	2	2	
Comprehension/Critical Literacy	21	21	42%
Literal Understanding (3.1)	4	4	
Inferences and Interpretation (3.2)	4-6	6	
Summary and Generalization (3.3)	5-7	5	
Analysis and Evaluation (3.4)	6-8	6	
Literature	15	15	30%
Literary Genre (4.1)	4	4	
Literary Elements (4.2)	5-7	5	
Figurative Language/Sound Devices (4.3)	4-6	6	
Research and Information	8	8	16%
Accessing Information (5.1)	4	4	
Interpreting Information (5.2)	4	4	
Total Test	50	50	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.H
2010 PASS Blueprint and Actual Item Counts: Grade 3 Mathematics

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Patterns and Algebraic Reasoning	8	8	18%
Algebra Patterns (1.1)	4	4	
Problem Solving (1.2)	4	4	
Number Sense	7	7	16%
Place Value (2.1)	3-4	4	
Whole Numbers and Fractions (2.2)	3-4	3	
Number Operations and Computation	12	12	27%
Estimation (3.1)	4	4	
Multiplication (3.2)	4	4	
Money Problems (3.3)	4	4	
Geometry and Measurement	12	12	27%
Spatial Reasoning (4.1)	4	4	
Measurement (4.2)	4	4	
Time and Temperature (4.4)	4	4	
Data Analysis and Probability	6	6	13%
Data Analysis (5.1)	2-4	4	
Probability (5.2)	2-4	2	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.I
2010 PASS Blueprint and Actual Item Counts: Grade 4 Mathematics

PASS Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Patterns and Algebraic Reasoning	8	8	18%
Algebra Patterns (1.1)	4	4	
Functions (1.2)	4	4	
Number Sense	10	10	22%
Place Value (2.1)	4	4	
Whole Numbers and Decimals (2.2)	2-4	3	
Fractions (2.3)	2-4	3	
Number Operations and Computation	11	11	24%
Multiplication (3.1)	2-4	3	
Division (3.2)	2-4	4	
Estimation (3.3)	4-5	4	
Geometry and Measurement	10	10	22%
Lines and Angles (4.1)	2-4	4	
Spatial Reasoning (4.3)	2-4	2	
Measurement (4.4)	4	4	
Data Analysis and Probability	6	6	13%
Data Analysis (5.1)	6	6	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.J
2010 PASS Blueprint and Actual Item Counts: Grade 5 Mathematics

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Patterns and Algebraic Reasoning	8	8	18%
Algebra Patterns (1.1)	4	4	
Problem Solving (1.2)	4	4	
Number Sense	8	8	18%
Fractions/Decimals/Percents (2.1)	4	4	
Number Theory (2.2)	4	4	
Number Operations and Computation	8	8	18%
Estimation (3.1)	4	4	
Whole Numbers/Decimals/Fractions (3.2)	4	4	
Geometry and Measurement	12	12	27%
Geometric Figure Properties (4.1)	4	4	
Perimeter/Area (4.2)	4	4	
Convert Measurements (4.5)	4	4	
Data Analysis and Probability	9	9	20%
Data Analysis (5.1)	5	5	
Probability (5.2)	4	4	
<i>Total Test</i>	<i>45</i>	<i>45</i>	<i>100%</i>

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.K
2010 PASS Blueprint and Actual Item Counts: Grade 6 Mathematics

PASS Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Algebraic Reasoning	10	10	22%
Patterns (1.1)	5	5	
Order of Operations (1.2)	5	5	
Number Sense	13	13	29%
Multiply/Divide Fractions (2.1)	2-3	2	
Decimals (2.2)	2-3	3	
Estimation (2.3)	4	4	
Expressions (2.5)	4	4	
Geometry	6	6	13%
Angles (3.1)	2-4	4	
Congruent and Similar Figures (3.2)	2-4	2	
Measurement	7	7	16%
Compare/Convert Units (4.2)	3-4	3	
Estimate Measurements (4.3)	3-4	4	
Data Analysis and Statistics	9	9	20%
Collect/Organize/Interpret Data (5.1)	2-3	2	
Construct/Interpret Graphs (5.2)	2-3	3	
Median/Mode (5.3)	4	4	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.L
2010 PASS Blueprint and Actual Item Counts: Grade 7 Mathematics

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Algebraic Reasoning	8	8	18%
Properties (1.1)	4	4	
Linear Equations (1.2)	4	4	
Number Sense	12	12	27%
Integers (2.1)	4	4	
Ratio/Proportion/Percent (2.2)	4	4	
Exponents (2.3)	4	4	
Geometry	9	9	20%
Geometric Figures (3.1)	2-3	3	
Angles (3.2)	2-3	2	
Coordinate System (3.3)	4	4	
Measurement	7	7	16%
Area and Perimeter (4.1)	2-4	3	
Customary/Metric Measurements (4.2)	2-4	4	
Data Analysis and Probability	9	9	20%
Outcomes/Simple Probability (5.1)	4	4	
Probability with Or, And, or Not (5.2)	2-3	3	
Combinations/Permutations (5.3)	2-3	2	
<i>Total Test</i>	<i>45</i>	<i>45</i>	<i>100%</i>

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.M
2010 PASS Blueprint and Actual Item Counts: Grade 8 Mathematics

PASS Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Algebraic Reasoning	9	9	20%
Equations (1.1)	5	5	
Inequalities (1.2)	4	4	
Number Sense	8	8	18%
Rational Numbers/Proportions (2.1)	4	4	
Exponents (2.2)	4	4	
Geometry	8	8	18%
Classify Solids (3.1)	4	4	
Pythagorean Theorem (3.2)	4	4	
Measurement	12	12	27%
Estimate Surface Area/Volume (4.1)	4	4	
Similar Figures (4.2)	4	4	
Formulas (4.3)	4	4	
Data Analysis and Statistics	8	8	18%
Data Representation (5.1)	4	4	
Central Tendency (5.2)	4	4	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.N.1
2010 PASS Blueprint and Actual Item Counts: Grade 5 Science Process Standards

PASS Process Standards & Objectives	Ideal Number of Items for Alignment to PASS*	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Observe and Measure	10	10	22%
SI Metric (P1.1)	5	5	
Similar/different characteristics (P1.2)	5	5	
Classify	10	10	22%
Observable properties (P2.1)	5	5	
Serial order (P2.2)	5	5	
Experiment	11	11	24%
Experimental design (P3.2)	7	7	
Hazards/practice safety (P3.4)	4	4	
Interpret and Communicate	14	14	31%
Data tables/line/bar/trend and circle graphs (P4.2)	6	6	
Prediction based on data (P4.3)	4	4	
Explanations based on data (P4.4)	4	4	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.N.2
2010 PASS Blueprint and Actual Item Counts: Grade 5 Science Content Standards

<i>PASS</i> Content Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i>	Actual Number of Items on 2010 Test	Ideal Percentage of Items
Properties of Matter and Energy	18	18	44%
Matter has physical properties (1.1)	6	6	
Physical properties can be measured (1.2)	6	6	
Energy can be transferred (1.3)	6	6	
Organisms and Environments	12	12	29%
Dependence upon community (2.1)	6	6	
Individual organism and species survival (2.2)	6	6	
Structures of the Earth and the Solar System	11	11	27%
Weather patterns (3.2)	6	6	
Earth as a planet (3.3)	5	5	
Total Test	41*	41*	100%**

* Safety items are not included within the content blueprint

** The ideal percents are based on the total number of items on a test that are matched to the content standards and do not include items added for safety.

Table 1.3.O.1
2010 PASS Blueprint and Actual Item Counts: Grade 8 Science Process Standards

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Observe and Measure	8	8	18%
Qualitative/quantitative observations/changes (P1.1)	4	4	
SI (metrics) units/appropriate tools (P1.2 and P1.3)	4	4	
Classify	8	8	18%
Classification system (P2.1)	4	4	
Properties ordered (P2.2)	4	4	
Experiment	16	16	36%
Experimental design (P3.2)	6	6	
Identify variables (P3.3)	6	6	
Hazards/practice safety (P3.6)	4	4	
Interpret and Communicate	13	13	29%
Data tables/line/bar/trend and circle graphs (P4.2)	7	7	
Explanations/prediction (P4.3)	6	6	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.O.2
2010 PASS Blueprint and Actual Item Counts: Grade 8 Science Content Standards

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i>	Actual Number of Items on 2010 Test	Ideal Percentage of Items
Properties and Chemical Changes in Matter	7-8	8	19%
Chemical reactions (1.1)	3-4	4	
Conservation of matter (1.2)	3-4	4	
Motion and Forces	8	8	20%
Motion of an object (2.1)	4	4	
Object subjected to a force (2.2)	4	4	
Diversity and Adaptations of Organisms	9	9	22%
Classification (3.1)	5	5	
Internal and external structures (3.2)	4	4	
<i>Structures/Forces of the Earth/Solar System</i>	8	8	20%
Landforms result from constructive and destructive forces (4.1)	4	4	
Rock cycle (4.2)	4	4	
Earth's History	7-8	8	19%
Catastrophic events (5.1)	3-4	4	
Fossil evidence (5.2)	3-4	4	
<i>Total Test</i>	<i>41*</i>	<i>41</i>	<i>100%**</i>

* Safety items are not included within the content blueprint

** The ideal percents are based on the total number of items on a test that are matched to the content standards and do not include items added for safety.

Table 1.3.P
2010 PASS Blueprint and Actual Item Counts: Grade 5 Social Studies

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Early Exploration	8	8	13%
Expeditions (2.1)	4	4	
Native American Reaction (2.2)	4	4	
Colonial America	12	12	20%
Settlements and Migration (3.1)	4	4	
Colonial Life (3.2)	4	4	
Individuals and Groups (3.3)	4	4	
American Revolution	12	12	20%
Causes and Results (4.1)	4	4	
Declaration of Independence (4.3)	4	4	
Individuals (4.4)	4	4	
Early Federal Period	8	8	13%
Constitutional Provisions (5.2)	4	4	
Ratification and Rights (5.3)	4	4	
Geographic Skills	20	20	33%
Maps/Charts/Graphs Usage (7.1)	7	7	
Human/Environment Interaction (7.2)	5	5	
Historical Places (7.3)	4	4	
Westward Movement (7.4)	4	4	
Total Test	60	60	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.Q
2010 PASS Blueprint and Actual Item Counts: Grade 7 Geography

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Geographic Tools	4	4	9%
Map Concepts (1.2)	4	4	
Regions	12	12	27%
Regional Characteristics (2.1)	4	4	
Conflict/Cooperation (2.2)	4	4	
Locations (2.4)	4	4	
Physical Systems	8	8	18%
Climate/Weather (3.2)	4	4	
Natural Disasters (3.3)	4	4	
Human Systems	8	8	18%
World Cultures (4.1)	4	4	
Population Issues (4.5)	4	4	
Human/Environment Interaction	8	8	18%
Natural Resources (5.1)	4	4	
Human Modification (5.2)	4	4	
Geography Skills	5	5	11%
Maps/Charts/Graphs (6.1)	5	5	
Total Test	45	45	100%

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

Table 1.3.R
2010 PASS Blueprint and Actual Item Counts: Grade 8 U.S. History

<i>PASS</i> Standards & Objectives	Ideal Number of Items for Alignment to <i>PASS</i> *	Actual Number of Items on 2010 Test	Ideal Percentage of Items **
Social Studies Process Skills (1.0)	6	6	13%
Causes of the American Revolution (3.0)	5	5	11%
Results of the American Revolution (4.0)	5	5	11%
Governing Documents/Early Federal Period (5.0)	6	6	13%
Northern/Southern Economic Growth (6.0)	4	4	9%
Jacksonian Era (7.0)	4	4	9%
Cultural Growth and Reform (8.0)	4	4	9%
Westward Movement (9.0)	6	6	13%
Eve of War (10.0)	5	5	11%
<i>Total Test</i>	<i>45</i>	<i>45</i>	<i>100%</i>

* A minimum of 4 items is required to report results for an objective, and a minimum of 6 items is required to report a standard. While the actual numbers of items on the test may not match the blueprint exactly, each future test will move toward closer alignment with the ideal blueprint.

** Percents are approximations and may result in a sum other than 100 due to rounding.

CHAPTER II. STATISTICAL ANALYSIS

This section provides an overview of the research and statistical analyses carried out for the 2010 administration of the OCCT. Following the administration of the OCCT, student demographic and item response data were transmitted to DRC's Psychometric Services (PS) department. PS staff is responsible for analyzing the OCCT test data and producing the scoring tables used for reporting.

The analyses of the test data can be broken down into several components: 1) classical item analyses; 2) differential item functioning (DIF) analyses; 3) reliability analyses; 4) calibration and equating; 5) production of scoring tables; and 6) validity analyses. In the following sections, the analysis procedures for each component are described. Separate sections are provided for the Multiple-Choice tests and the Writing tests.

2.1 Data Files for Statistical Analysis

Preliminary item and DIF analyses and the final calibration/equating for the Multiple-Choice tests were conducted using early-return sample data consisting of approximately 50% of the examinee population. The final item and DIF analyses were conducted using data files that contained 100% of the student data. For the analysis of the Writing tests, the rater-year effect analysis was conducted based on a sample ($n = 510$) which was randomly drawn from 2007 and rescored in 2010. All other analyses on writing were conducted using the final population data file.

2.2 Analysis of the Multiple-Choice Tests

2.2.1 Classical Item Analyses

Classical item analyses were conducted using DRC's software, iTEMs (DRC, 2010). The analyses involved computing a set of statistics based on classical test theory for every item in each form. Each statistic is designed to provide empirical information about the characteristics of each item. The statistics estimated for OCCT items are described below.

Classical Item Difficulty (“ p -value”):

This statistic indicates the proportion of examinees in the sample that answered the item correctly. Desired p -values generally fall within the range of 0.30 to 0.90. Occasionally, items that fall outside this range can be justified for inclusion in an assessment based upon other quality indicators (e.g., adequate point-biserial), the educational importance of the item's content, or to better measure students with very high or low achievement.

Item Discrimination (“point-biserial”):

This statistic describes the relationship between performance on the specific item and performance on the entire form. Estimated as the correlation between the item score and the total test score, it indicates the extent to which test takers with high test scores tend to answer the item correctly, and those with low scores tend to answer incorrectly. The point-biserial correlation for item i is given by

$$pbis_i = \frac{\mu_+ - \mu_x}{\sigma_x} \sqrt{\frac{p_i}{1 - p_i}},$$

where μ_+ is the mean score for those students who answered item i correctly, and μ_x and σ_x are the mean score and standard deviation for the test form, and p_i is the item difficulty for item i .

Items with negative or low correlations can indicate problems with the item (e.g., incorrect key, multiple correct answers or unusually complex content), or can indicate that students have not been taught the content.

Examination of Empirical Item Response Curves (EIRC):

iTEMs provides graphical displays of student performance on each item. In the MC item plots, the x-axis represents the criterion score level (the total number-correct score) and y-axis represents the percentage of examinees choosing the response option. Each response option is plotted, representing the percentage of examinees that chose that particular option by ability level. One would expect the curve for the correct option to increase as ability level increases. These graphs were reviewed by DRC psychometricians.

Percentage of Students Omitting or not Reaching an Item:

This statistic is useful for identifying issues related to testing time and item/test layout. Testing time issues do not exist for the OCCT's as they are untimed. However, if the omit percentage is greater than 5% for a single item, this could be an indication of an item/test layout problem. For example, students might accidentally skip an item that follows a lengthy stem.

For the OCCT operational and field-test analyses, a series of flags were created in order to automatically identify items with performance characteristics that are at times considered unusual. The following flagging criteria were applied to all items tested in spring 2010:

- P -value less than .30 or greater than .90;
- The percentage choosing an “incorrect” option is equal to or greater than the percentage choosing the correct option;
- The percentage of students selecting any of the “incorrect” options is larger than 30%;
- Point-biserial correlation is less than .30 for the correct answer;
- Any of the “incorrect” answer options (distractors) with a positive point-biserial;
- Percentage of test takers omitting the item is greater than 5%.

After the operational administration, the early-return data were used to conduct preliminary analyses to verify the accuracy of the scoring keys and to obtain an early indication of how items were functioning. Content specialists examined all flagged items to ensure that the items were correctly keyed.

Upon receipt of the complete student data, the items were further scrutinized during a final round of classical item analysis. After content specialist's review and verification, item statistics were prepared for uploading to the item bank.

Summary statistics describing the difficulty and discrimination of items comprising the operational forms are given in Table 2.2.A. Results are combined across test forms for a given grade and content area because the operational items were the same for all test forms. Differential item functioning (DIF) flags and reliability indices (alpha, SEM, and stratified alpha) are also provided. These statistics are described in the sections that follow.

Table 2.2.A
Classical Item and Test Analyses Summary for the Operational Forms

Content Area/Grade	Grade	N Items	Alpha	SEM	Stratified Alpha	<i>p</i> -value		Pt-Biserial		Flagged Item Count	
						Mean	SD	Mean	SD	Stats ^a	DIF
Reading	3	50	0.89	2.90	0.89	0.69	0.14	0.40	0.07	10	0
	4	50	0.89	2.83	0.89	0.71	0.14	0.40	0.07	7	0
	5	50	0.89	2.70	0.89	0.74	0.14	0.41	0.07	10	2
	6	50	0.90	2.99	0.90	0.66	0.12	0.41	0.08	8	2
	7	50	0.87	2.88	0.87	0.70	0.15	0.37	0.07	15	2
	8	50	0.88	2.83	0.88	0.72	0.15	0.38	0.08	15	5
Mathematics	3	45	0.90	2.58	0.91	0.74	0.13	0.43	0.08	5	0
	4	45	0.89	2.62	0.89	0.72	0.15	0.42	0.07	8	0
	5	45	0.89	2.77	0.89	0.67	0.14	0.42	0.08	8	2
	6	45	0.89	2.81	0.90	0.67	0.11	0.42	0.08	5	0
	7	45	0.87	2.82	0.87	0.65	0.16	0.39	0.06	9	0
	8	45	0.88	2.82	0.88	0.68	0.11	0.40	0.07	9	1
Science	5	45	0.88	2.72	0.88	0.69	0.15	0.40	0.06	11	0
	8	45	0.86	2.84	0.86	0.65	0.16	0.38	0.07	14	2
Social Studies	5	60	0.89	3.41	0.89	0.60	0.14	0.37	0.08	20	0
Geography	7	45	0.86	2.90	0.86	0.64	0.13	0.38	0.08	16	4
U.S. History	8	45	0.89	2.88	0.89	0.62	0.13	0.41	0.07	9	0

^aClassical item statistics flagged using the criteria from the bulleted points above.

As shown in Table 2.2.A, the mean item difficulties of the tests across content/grade ranged from 0.60 to 0.74, and the mean point-biserial correlations ranged from 0.37 to 0.43. The internal consistency reliability estimates (coefficient alpha) of all tests were high, ranging from 0.86 to 0.91. The stratified alpha coefficients were almost identical to the alpha (after rounding). The SEMs ranged from 2.58 to 3.41.

Table 2.2.A also shows that a small number of operational test items were flagged for out-of-range statistics and/or C category DIF. Items flagged for out-of-range statistics were scrutinized by content experts to verify the accuracy of the items in the test books, to verify keys, and to judge whether items were performing as expected. All items flagged for out-of-range statistics were found to be accurate, correctly keyed, and performing in a satisfactory manner with respect to content. DIF results for the operational items are discussed in the next section.

The results of the classical item analysis for the field-test items are presented in Table 2.2.B. Field-test items with extreme difficulty values, low point-biserials or poorly functioning distractors, and/or DIF were flagged for review by content experts. The number of items flagged for poor statistics ranged from 34 to 59 per content/grade. A very small number of items were

flagged for DIF. All flagged items were evaluated at a Data Review Meeting and individual decisions were made about each item. Items rejected during Data Review will not be eligible as operational items in the future test administrations. Items accepted with revisions will be returned to the item bank and will be revised and re-field tested as necessary. The results of the Data Review are presented in Appendix A.

Table 2.2.B
Item Analyses Summary for Field-test Items

Content Area	Grade	Sample Size	N Items	p-value		Pt-Biserial		Flagged Item Count	
		Range		Mean	SD	Mean	SD	Stats ^a	DIF
Reading	3	5323-5535	80	0.61	0.17	0.36	0.10	30	2
	4	4800-5452	80	0.62	0.16	0.34	0.10	28	3
	5	4259-5493	80	0.62	0.17	0.34	0.12	27	0
	6	4692-5398	80	0.57	0.17	0.32	0.12	29	1
	7	4176-5255	80	0.65	0.16	0.34	0.11	28	1
Mathematics	8	4939-5526	80	0.65	0.16	0.30	0.11	39	5
	3	5435-5603	80	0.75	0.16	0.36	0.10	31	3
	4	4953-5511	80	0.64	0.16	0.36	0.10	32	1
	5	4407-5515	80	0.58	0.23	0.33	0.13	45	2
	6	4766-5410	80	0.55	0.20	0.31	0.13	36	1
Science	7	4198-5267	80	0.54	0.20	0.30	0.10	45	2
	8	4844-6058	80	0.59	0.16	0.33	0.11	32	1
	5	4453-5574	80	0.56	0.18	0.30	0.11	36	1
	8	4375-5220	80	0.55	0.19	0.30	0.11	40	3
	5	4985-5952	80	0.47	0.14	0.29	0.12	45	0
Geography	7	5157-8292	80	0.53	0.17	0.27	0.13	47	1
History	8	4798-5515	80	0.50	0.15	0.33	0.13	30	2

^aOut-of-range classical item statistics

2.2.2 Differential Item Functioning Analyses

One of the goals of test development is to assemble a set of items that provides an estimate of a student's ability that is as fair and accurate as possible for all groups within the population. DIF statistics are used to identify items that groups of students with the same underlying level of ability have different probabilities of answering correctly. If the item is differentially more difficult for an identifiable subgroup when conditioned on ability, the item may be measuring something different from the intended construct. However, it is important to recognize that DIF-flagged items might be related to actual differences in relevant knowledge or skills (item impact) or a statistical Type I error (a "false positive"). As a result, DIF statistics are used to identify potential sources of item bias. Subsequent review by content experts and bias/sensitivity committees is required to determine the source and meaning of performance differences.

For the OCCT, the Mantel-Haenszel (MH) procedure (Mantel & Haenszel, 1959; Holland & Thayer, 1988) was used to estimate DIF statistics for subgroups of interest defined by the SDE for NCLB accountability. Comparison groups were based on gender (female versus male), ethnicity (Hispanic versus White, American Indian versus White, African American versus White, Asian versus White, Pacific Islander versus White), and economic status (students who are economically disadvantaged as indicated by participation in a free and reduced-price school lunch program versus students who are not economically disadvantaged). Items with statistically significant differences in performance were flagged for possible biased or unfair content that was undetected in earlier fairness and bias reviews. DIF analyses results were not considered as valid

when the sample size for either the reference group (i.e., male, White, not economically disadvantaged) or focal group (i.e., female, Hispanic, American Indian, African American, Asian, Pacific Islander, economically disadvantaged) was less than 300 *and* the sample size for the two groups combined was less than 700.

The MH procedure is one of the more commonly used methods to detect DIF. This method uses contingency tables to compare the probability of success on each item for the studied groups of interest after matching on overall ability (i.e., total test score). The common odds ratio is estimated across all categories of matched examinee ability. The resulting estimate is interpreted as the relative likelihood of success on a particular item for members of two groups when matched on ability. As such, the common odds ratio provides an estimated effect size where a value of unity indicates equal odds, and thus no DIF (Dorans & Holland, 1993).

The common odds ratio (α) is estimated as $\hat{\alpha}_{MH} = \frac{\sum_{s=1}^S R_{rs} W_{fs} / N_{ts}}{\sum_{s=1}^S R_{fs} W_{rs} / N_{ts}}$, where

R_{rs} = the number of examinees in the reference group who answer the item correctly,

W_{fs} = the number of examinees in the focal group who answer the item incorrectly,

R_{fs} = the number of examinees in the focal group who answer the item correctly,

W_{rs} = the number of examinees in the reference group who answer the item incorrectly,

N_{ts} = the total number of examinees.

The odds ratio takes on values from 0 to infinity and is interpreted as the average factor by which the odds that an examinee of the reference group will answer an item correctly exceed that of a member of the comparable focal group. The statistical test is $H_0: \alpha = 1$, where α is a common odds ratio assumed equal for all matched score categories $s = 1$ to S . Values less than unity indicate DIF in favor of the focal group, a value of unity indicates the null condition, and a value greater than one indicates DIF in favor of the reference group. The associated MH χ^2 is distributed as a chi-square random variable with 1 degree of freedom. As an index of magnitude, the odds ratio is frequently transformed to a delta scale given by $MH\ D-DIF = -2.35 \ln(\hat{\alpha}_{MH})$ where negative values indicate DIF in favor of the reference group and positive values favor the focal group.

A classification scheme puts items into three DIF categories on the basis of a combination of statistical significance and magnitude (absolute value) of MH D-DIF (Zwick and Ercikan, 1989):

A-items or negligible DIF: MH D-DIF is not statistically different from 0 (at the .05 level) or its absolute value is less than 1 delta unit;

B-items or intermediate DIF: MH D-DIF is statistically different from 0 (at the .05 level) and its absolute value is at least 1 but less than 1.5 or an absolute value of at least 1 but not significantly greater than 1 (at the .05 level);

C-items or large DIF: MH D-DIF is statistically different from 1 (at the .05 level) and its absolute value is at least 1.5.

Items classified as B+ or C+ tend to be easier for members of the focal group than for members of the reference group whose total scores on the test are like those of the focal group. Items classified as B- or C- tend to be harder for members of the focal group than for members of the reference group whose total scores on the test are like those of the focal group.

Items classified in category C were sent to test development staff for review. They were asked to consider any identifiable characteristics that may have contributed to the differential item functioning. The items were then submitted to the SDE for further review.

Table 2.2.A shows that a small number of operational items were flagged for C DIF. These items were reviewed by DRC's content experts. Recommendations were made by DRC on whether to remove an item with C DIF from scoring or not. SDE content experts further reviewed these items and made the final decision. As a result, no items were dropped in the 2010 administration.

DIF analysis was also conducted on the field tests. Items with C DIF were flagged and reviewed by SDE and DRC's content experts at the data review meeting. Appendix A reports the items rejected due to DIF and/or other poor statistics.

2.2.3 Item Calibration and Equating

The purpose of item calibration and equating is to create a common scale for expressing the difficulty estimates of all the items across forms within a test. The scale is initially defined so that the examinees used in the calibration will have a mean score of 0 and a standard deviation of 1. It should be noted that the metric of this scale is often referred to as the "theta" metric. This scale is not used for reporting purposes because its values typically range from -3.0 to +3.0, which is a scale that is not easily understood. Therefore, following calibration and equating, the scale is usually transformed to a reporting scale that can be understood more easily by students, teachers, and other stakeholders.

The three-parameter logistic (3PL) model was used to calibrate the OCCT test items. The 3PL model expresses the probability that a person with ability θ will respond correctly to item j as a function of item and ability parameters:

$$P(U_j = 1 | \theta) = P_j(\theta) = c_j + \frac{1 - c_j}{1 + e^{-1.7a(\theta - b)}},$$

where:

U_j is the response to item j , 1 if correct and 0 if incorrect;

a_j is the slope parameter of item j , characterizing its discriminating power;

b_j is the threshold parameter of item j , characterizing its difficulty; and

c_j is the lower asymptote parameter of item j , reflecting the chance that students with very low proficiency will select the correct answer, sometimes called the "pseudo-guessing" level.

The parameters estimated for the 3PL model were discrimination (a), difficulty (b), and the pseudo-guessing level (c). All item response theory (IRT) based analyses were conducted using PARSCALE (Muraki and Bock, 2003).

For each operational test, items were calibrated separately by content and grade. The calibrations were examined to assess the quality of the parameter estimates and model-data fit. Items were flagged for:

- a -parameters less than 0.3 or greater than 2.3
- b -parameters less than -3.5 or greater than 3.5
- c -parameters greater than 0.35 for 4-option items
- Not calibrated due to biserial correlations less than 0.10
- Data model fit is “bad.” (This criterion varies depending on the response n-count.)

Flagged items were reviewed to determine whether they should be excluded from scoring. No items in the 2010 OCCT were excluded from scoring because of IRT results.

After the final set of item parameter estimates were established, the scales for the 2010 operational tests were linked to the reporting scale using the test characteristic curve (TCC) method described by Stocking and Lord (1983). The Stocking and Lord procedure involves finding a linear transformation that will minimize the sum of squared differences between two TCCs generated from two sets of anchor item parameters.

Embedded in the 2010 OCCT were sets of anchor items that had served as operational items in the 2009 OCCT. These items were positioned so their sequences were very similar to that in the prior year. The sets were chosen for being both content and statistically representative of the entire test to ensure an accurate equating result. The anchor set is mostly unique for each testing cycle, though some items may be used as anchors for multiple administrations. Repeated use of an item creates the risk of overexposure and is avoided in practice. Table 2.2.C summarizes the number of anchor items per test.

The parameters for the 2009 items were expressed on the reporting scale. These 2009 item parameters served as a reference item set and were used with their 2010 counterparts and the Stocking and Lord procedure to find transformation constants. These constants were used to transform the 2010 item parameters so that they were expressed on the reporting scale. Once this was done, the transformed parameters were used to generate raw score to scaled score conversion tables.

Table 2.2.C
Number of Anchor Items by Subject and Grade

Subject	Grade	Number of Operational Items per Form ^a	Number of Operational Items in Anchor Set ^a
Reading	3-8	50	17, 18, 16, 14, 17, and 15 respectively
Writing	5, 8	1	0
Mathematics	3-8	45	20, 19, 18, 18, 17, and 17 respectively
Science	5, 8	45	16,14
Social Studies	5	60	20
Geography	7	45	15
U.S. History	8	45	15

^a Operational item counts include anchor items.

Field-test items were placed on the operational scale in a similar fashion. For each content/grade, field-test items were calibrated with the operational items. Resulting field-test parameters were placed on the OCCT reporting scale using the operational items as the anchor set in the Stocking and Lord procedure.

2.2.4 Raw Score to Scaled Score Conversion

Since 2005, the OCCT scaled scores have been produced using a number-correct scoring procedure that is based on IRT. This procedure produces maximum-likelihood trait estimates for each obtainable raw score, except for raw scores at chance or below-chance levels and the perfect raw score. It is conventional to assign scaled scores to at and below-chance level raw scores and perfect raw scores using a rational, but not necessarily maximum likelihood, procedure. These values are called the lowest obtainable scaled score (LOSS) and the highest obtainable scaled score (HOSS). The LOSS and HOSS values assigned to all OCCT operational tests were 400 and 990, respectively.

For all MC tests, the OCCT score scale uses a three-digit integer that spans a range from 400 (LOSS) to 990 (HOSS). The Proficient cut score for reading and Mathematics and the Satisfactory cut score for science and social studies is 700 for all tests. The raw-score to scaled-score conversion tables are provided in Appendix B.

2.2.5 Test Score Reliability

Test score reliability focuses on the extent to which differences in test scores reflect true differences in the knowledge, ability, or skills being tested rather than random fluctuations. The variance in the distributions of test scores, essentially the differences among individuals, is partly due to real differences in the knowledge, skills, or ability being tested (called true score variance) and partly due to random factors that cause variability in examinee performance (called error variance). The number used to describe reliability is an estimate of the proportion of true score variance to total variance. Several different ways of estimating this proportion exist.

Coefficient Alpha

When the goal is to estimate the precision of a set of test scores from a single administration, a measure of internal consistency (sensitive to random errors associated with item content sampling) is frequently used to estimate reliability. For the OCCT, the measure of internal

consistency called coefficient alpha (Cronbach, 1951) was used to estimate the reliability of the test scores. The formula for coefficient alpha is given by

$$\rho_{xx'} \geq \frac{k}{k-1} \left(1 - \frac{\sum \sigma_i^2}{\sigma_x^2}\right),$$

where k is the number of items on the test, $\sum \sigma_i^2$ is item score variance summed over all items, and σ_x^2 is observed-score variance.

Internal consistency measures apply only to the test form being analyzed. They do not take into account form-to-form variation due to equating limitations, nor are they sensitive to day-to-day variation due, for example, to state of health or testing environment. Reliability coefficients may range from 0 to 1. The higher the reliability coefficient for a set of scores, the more likely it would be for individuals to obtain very similar scores over replicated testing (e.g., using the same number of items, sampling same content domain(s), etc.). The internal consistency of the multiple-choice test scores are reported in Table 2.2.A for all examinees and in Tables 3.3 – 3.21 by demographic subgroup.

When a test contains different components (e.g., content standards), the stratified alpha coefficient can provide a more accurate estimate of the overall test reliability (Qualls, 1995). The stratified alpha coefficient is calculated by

$$\text{Strat } \alpha \rho_{xx'} = 1 - \frac{\sum \sigma_{x_j}^2 (1 - \alpha \rho_{x_j x'_j})}{\sigma_{Total}^2},$$

where, σ_{Total}^2 is the variance of the total test scores; $\sigma_{x_j}^2$ is the variance of scores for each test component (i.e., content standards in this case); and $\alpha \rho_{x_j x'_j}$ is the coefficient alpha reliability for scores from content standard J . The stratified alpha coefficients for the multiple-choice test scores are reported in Table 2.2.A for all examinees and in Tables 3.3-3.21 by demographic subgroup.

Standard Error of Measurement

The standard error of measurement (SEM) is the standard deviation of the errors of measurement that is associated with the test scores of a specific group of test takers. In Classical Test theory (CTT), an overall SEM can be estimated as a function of the standard deviation of observed scores and test reliability coefficient:

$$SEM = s_x \sqrt{1 - r_{xx'}},$$

where SEM is standard error of measurement, s_x is standard deviation of observed scores, and $r_{xx'}$ is a coefficient of reliability.

The SEM is particularly useful in determining the confidence interval (CI) that captures an examinee's true score. Assuming that measurement error is normally distributed, it can be said that upon infinite testing replications, approximately 95 percent of the CIs of ± 1.96 SEM

around the observed score would contain an examinee's true score (Crocker & Algina, 1986). For example, if an examinee's observed score on a given test equals 15 points, and SEM equals 1.92, one can be 95% confident that the examinee's true score lies between 11 and 19 points (15 ± 3.76 rounded to the nearest integer). Table 2.2.A provides the SEM for each multiple-choice test.

Conditional Standard Error of Measurement

From the IRT framework, a standard error of measurement can be estimated for each measured ability. Thus, it is often referred to as a conditional standard error of measurement (CSEM). The expected a posteriori estimation of CSEM proposed by Kolen, Zeng, and Hanson (1996) was used for the OCCT. The calculation of CSEM can be expressed as:

$$CSEM(S_X | \theta) = \sqrt{\left[\sum_{X=0}^{MaxX} S_X^2 p(X | \theta) \right] - \left[\sum_{X=0}^{MaxX} S_X \cdot p(X | \theta) \right]^2}$$

where S_X is the scaled score for a particular number correct score X ; θ is the IRT ability scaled value conditioned on; and $p(X | \theta)$ is the probability function that is computed using a recursive algorithm given by Thissen, Pommerich, Billeaud, and Williams (1995). For the operational OCCT, CSEMs were provided for each obtainable scaled score (see Appendix B).

Reliability of Performance-Level Classification Decisions

Student performance on the OCCT is classified into one of four achievement levels using cut scores adopted by the SDE. Table 2.2.D provides the cut score for each achievement level and the CSEM associated with each cut score in 2010.

Table 2.2.D
Conditional Standard Errors of Measurement for Each Achievement Level Cut Score

Content Area	Grade	Limited Knowledge		Satisfactory		Advanced	
		Scale Score	CSEM	Scale Score	CSEM	Scale Score	CSEM
Reading	3	649	26	700	23	891	50
	4	658	21	700	21	845	49
	5	641	23	700	23	830	47
	6	647	26	700	21	828	39
	7	668	24	700	22	802	32
	8	655	26	700	25	833	41
Mathematics	3	636	28	700	23	798	36
	4	639	29	700	25	816	38
	5	642	32	700	23	767	25
	6	660	25	700	20	754	22
	7	667	28	700	26	766	24
	8	662	27	700	24	771	27
Science	5	638	49	700	26	814	24
	8	647	62	700	28	829	24
Social Studies	5	645	37	700	24	786	20
Geography	7	595	70	700	33	847	33
History	8	622	49	700	25	821	32
Writing ^a	5	26	na	36	na	54	na
	8	25	na	36	na	54	na

^aWriting cut scores are in the composite score metric.

The reliability of 2010 achievement-level classification decisions was assessed using the computer program BB-CLASS (Brennan, 2004), which provides two statistics that describe the reliability of classifications based on test scores (Livingston & Lewis, 1995). More specifically, information from an administration of one form is used to estimate the following:

Decision Accuracy, which describes the extent to which performance-level classification decisions based on the administered test form would agree with the decisions that would be made on the basis of a perfectly reliable test (i.e., meaning if it was possible to know each examinee's true score). Decision accuracy answers the question: How does the actual classification of test takers, based on their single-form scores, agree with the classification that would be made on the basis of their true scores, if their true scores were somehow known?

Decision Consistency, which describes the extent to which classification decisions based on the administered test form would agree with the decisions made if a parallel alternate form had been administered. Decision consistency answers the question: What is the agreement between the classifications based on two non-overlapping, equally difficult forms of the test?

For each performance level and test, true scores and single-form scores on forms parallel to the one actually given are estimated following the Livingston and Lewis (1995) method. The decision accuracy is estimated using an estimated joint distribution of reported performance level classifications on the current form of the exam and the performance-level classifications based on the true score. Decision consistency is estimated using an estimated joint distribution of

reported performance-level classifications on the current form of the exam and performance-level classifications on the parallel alternate form.

In each case, the proportion of performance-level classifications with exact agreement is the sum of the entries in the diagonal of the contingency table representing the joint distribution. Reliability of classification at each performance-level cut score is estimated by collapsing the joint distribution at the passing score boundary into a 2-by-2 table and summing the two entries.

Table 2.2.E provides the results for decision accuracy and consistency analyses that were conducted at the Limited Knowledge, Proficient/Satisfactory, and Advanced cut scores, and for the four performance levels (total). It should be noted that decision accuracy and consistency indices for the four performance levels should be lower than those for each cut, as shown in Table 2.2.D. This is not surprising since classification using four levels would allow more opportunity to change achievement levels. Hence there would be more classification errors in the four achievement levels, resulting in lower consistency indices.

For the OCCT, a *PASS*ing score is one that meets or exceeds the Proficient/Satisfactory cut score. Across all tests, the decision accuracy of the Proficient/Satisfactory cut scores ranged from 0.89–0.94 and decision consistency ranged from 0.85–0.92. These results indicate that at least 89% students meeting or exceeding the Proficient/Satisfactory cut score would receive the same *PASS*/fail classification if their true scores were known. If a parallel test were administered, at least 85% or more of students meeting or exceeding the Proficient/Satisfactory cut score would be classified in the same way.

Table 2.2.E
Estimates of the Reliability of Decisions for Specified Cut Scores ^a

Content Area	Grade	Decision Accuracy				Decision Consistency			
		Limited Knowledge	Proficient/Satisfactory	Advanced	Total	Limited Knowledge	Proficient/Satisfactory	Advanced	Total
Reading	3	0.95	0.91	0.98	0.84	0.93	0.88	0.97	0.78
	4	0.94	0.91	0.97	0.82	0.91	0.87	0.96	0.76
	5	0.95	0.91	0.91	0.77	0.93	0.87	0.90	0.71
	6	0.94	0.91	0.94	0.80	0.92	0.88	0.93	0.73
	7	0.93	0.90	0.91	0.75	0.90	0.86	0.88	0.67
	8	0.94	0.91	0.91	0.77	0.92	0.87	0.88	0.69
Mathematics	3	0.96	0.92	0.91	0.79	0.94	0.89	0.88	0.71
	4	0.95	0.91	0.92	0.78	0.92	0.87	0.89	0.70
	5	0.95	0.91	0.91	0.77	0.92	0.87	0.87	0.68
	6	0.93	0.91	0.91	0.76	0.90	0.87	0.87	0.67
	7	0.92	0.89	0.90	0.72	0.88	0.85	0.87	0.64
	8	0.93	0.90	0.91	0.75	0.90	0.86	0.87	0.66
Science	5	0.98	0.94	0.90	0.82	0.97	0.92	0.86	0.76
	8	0.98	0.94	0.92	0.83	0.96	0.91	0.88	0.77
Social Studies	5	0.94	0.91	0.92	0.78	0.92	0.88	0.89	0.70
Geography	7	0.98	0.92	0.91	0.81	0.97	0.89	0.87	0.74
History	8	0.95	0.91	0.93	0.79	0.93	0.88	0.91	0.72

a: The analysis was based on the final data files with students who took the standard OCCT.

2.2.6 Validity

As noted in the *Standards for Educational and Psychological Testing*, “validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of the tests” (AERA, APA, & NCME, 1999, P.9). Content representativeness

considerations, item bias (i.e., DIF) analysis, and correlations among content standards are often used as sources of validity evidence.

Each test's blueprint specifies the proportion of items that should be devoted to any given content unit. The blueprint is used as a guide by test developers when assembling a test from a pool of candidate items that are classified by content unit. Validity evidence related to test content is bolstered to the extent that the numbers of items allocated to each *PASS* Standard/Objective reflect what is specified by the test blueprints. Tables 1.2 to 1.18 in Chapter 1 provide content validity evidence by standard for the 2010 OCCT.

Differential item functioning with respect to gender, ethnicity, and economic status helps address construct-irrelevant variance, which represents an important threat to the validity of achievement tests. As noted in the section of Differential Item Functioning Analyses, field-test items are screened and reviewed for DIF by SDE content specialists. Only items approved by SDE are eligible for operational use. DIF analyses were also conducted on the operational items. After SDE and DRC's content specialists' review, no item was dropped from the operational tests. The number of operational and field-test items with C DIF were reported in Tables 2.2.A and 2.2.B.

Intercorrelations among standards provide evidence of convergent test validity. The analyses were performed by summing the obtained raw score points for each standard and then correlating the subtotals associated with each standard. Standards with low point totals, e.g., less than five, usually have markedly attenuated coefficients, meaning that they will be spuriously low in magnitude. Tables 1.3.B to 1.3.R list the numbers of items associated with each standard. The correlations among standards are reported in the left corner of Tables 2.2.F through 2.2.V.

The correlations corrected for attenuation are reported in the right corner of Tables 2.2.E through 2.2.U. Correcting for attenuation adjusts the correlation between the two measures to account for the unreliability of both. Although the theoretical upper bound for a correlation is 1.0, disattenuated correlations can be greater. This is often seen in practice when the correlations are relatively high and the reliabilities relatively low. However, two underlying factors should be noted. The first is that sample statistics are being used to estimate population parameters. The second, and likely more prevailing issue, is that something akin to a "design misspecification" occurs. The internal consistency reliability indices used for the OCCT likely do not capture all the sources of random error in the test scores, and, as such, might over estimate reliability. One might also postulate potential negative biases (e.g., lack of item homogeneity due to multidimensional content standards). Thus, it is possible that any given tabled disattenuated correlation may be too high, or too low, depending on which bias prevails. Also note that the correlations between standards and total test are spuriously inflated given they have items in common.

Given that none of these tests have perfect reliabilities (equal to one), the disattenuated correlations are somewhat higher than the correlations. Disattenuated correlations less than 1.0 suggest that the different strands are measuring slightly different aspects of the constructs. Values around 1.0 suggest that the same or very similar constructs are being measured.

Table 2.2.F
Standards Intercorrelation: Grade 3 Reading

	Reading	Vocabulary	Comprehension /Critical Literacy	Literature	Research and Information
Reading	--	1.09	1.12	1.10	1.06
Vocabulary	0.84	--	0.94	0.96	0.96
Comprehension/Critical Literacy	0.94	0.68	--	0.99	0.90
Literature	0.83	0.63	0.70	--	0.90
Research and Information	0.72	0.56	0.57	0.52	--

Table 2.2.G
Standards Intercorrelation: Grade 4 Reading

	Reading	Vocabulary	Comprehension /Critical Literacy	Literature	Research and Information
Reading	--	1.11	1.10	1.08	1.09
Vocabulary	0.82	--	0.95	0.88	0.97
Comprehension/Critical Literacy	0.94	0.67	--	0.92	0.94
Literature	0.80	0.54	0.65	--	0.94
Research and Information	0.74	0.55	0.61	0.53	--

Table 2.2.H
Standards Intercorrelation: Grade 5 Reading

	Reading	Vocabulary	Comprehension /Critical Literacy	Literature	Research and Information
Reading	--	1.09	1.13	1.13	1.11
Vocabulary	0.87	--	0.97	0.98	0.99
Comprehension/Critical Literacy	0.92	0.71	--	1.02	0.97
Literature	0.87	0.67	0.71	--	0.99
Research and Information	0.74	0.59	0.59	0.56	--

Table 2.2.I
Standards Intercorrelation: Grade 6 Reading

	Reading	Vocabulary	Comprehension /Critical Literacy	Literature	Research and Information
Reading	--	1.11	1.09	1.12	1.14
Vocabulary	0.79	--	0.97	0.99	1.01
Comprehension/Critical Literacy	0.93	0.66	--	0.99	0.98
Literature	0.88	0.62	0.74	--	1.02
Research and Information	0.77	0.54	0.64	0.60	--

Table 2.2.J
Standards Intercorrelation: Grade 7 Reading

	Reading	Vocabulary	Comprehension /Critical Literacy	Literature	Research and Information
Reading	--	1.12	1.13	1.16	1.15
Vocabulary	0.76	--	0.95	0.95	0.95
Comprehension/Critical Literacy	0.93	0.60	--	0.98	1.01
Literature	0.82	0.52	0.65	--	0.97
Research and Information	0.78	0.50	0.65	0.54	--

Table 2.2.K
Standards Intercorrelation: Grade 8 Reading

	Reading	Vocabulary	Comprehension /Critical Literacy	Literature	Research and Information
Reading	--	1.10	1.11	1.15	1.16
Vocabulary	0.67	--	0.97	0.92	0.98
Comprehension/Critical Literacy	0.93	0.56	--	0.97	1.01
Literature	0.87	0.48	0.70	--	1.00
Research and Information	0.77	0.45	0.63	0.57	--

Table 2.2.L
Standards Intercorrelation: Grade 3 Mathematics

	Mathematics	Patterns and Algebraic Reasoning	Number Sense	Number Operations and Computation	Geometry and Measurement	Data Analysis and Probability
Mathematics	--	1.05	1.07	1.04	1.11	1.03
Patterns and Algebraic Reasoning	0.82	--	0.94	0.84	0.93	0.90
Number Sense	0.78	0.59	--	0.85	0.98	0.94
Number Operations and Computation	0.86	0.60	0.57	--	0.89	0.84
Geometry and Measurement	0.85	0.61	0.60	0.62	--	0.94
Data Analysis and Probability	0.81	0.61	0.59	0.60	0.62	--

Table 2.2.M
Standards Intercorrelation: Grade 4 Mathematics

	Mathematics	Patterns and Algebraic Reasoning	Number Sense	Number Operations and Computation	Geometry and Measurement	Data Analysis and Probability
Mathematics	--	1.07	1.11	1.05	1.11	1.07
Patterns and Algebraic Reasoning	0.79	--	0.92	0.89	0.90	0.96
Number Sense	0.83	0.57	--	0.89	0.98	0.94
Number Operations and Computation	0.87	0.62	0.62	--	0.86	0.87
Geometry and Measurement	0.79	0.53	0.59	0.57	--	0.93
Data Analysis and Probability	0.76	0.56	0.56	0.57	0.52	--

Table 2.2.N
Standards Intercorrelation: Grade 5 Mathematics

	Mathematics	Patterns and Algebraic Reasoning	Number Sense	Number Operations and Computation	Geometry and Measurement	Data Analysis and Probability
Mathematics	--	1.11	1.14	1.09	1.10	1.03
Patterns and Algebraic Reasoning	0.80	--	1.01	0.94	0.95	0.89
Number Sense	0.80	0.58	--	0.99	0.97	0.93
Number Operations and Computation	0.82	0.58	0.59	--	0.94	0.84
Geometry and Measurement	0.87	0.61	0.61	0.63	--	0.88
Data Analysis and Probability	0.79	0.55	0.57	0.54	0.60	--

Table 2.2.O
Standards Intercorrelation: Grade 6 Mathematics

	Mathematics	Algebraic Reasoning	Number Sense	Geometry	Measurement	Data Analysis and Statistics
Mathematics	--	1.08	1.11	1.01	1.04	1.09
Algebraic Reasoning	0.84	--	0.96	0.85	0.86	0.91
Number Sense	0.89	0.67	--	0.84	0.93	0.94
Geometry	0.68	0.49	0.51	--	0.83	0.87
Measurement	0.81	0.58	0.65	0.48	--	0.88
Data Analysis and Statistics	0.80	0.58	0.62	0.47	0.56	--

Table 2.2.P
Standards Intercorrelation: Grade 7 Mathematics

	Mathematics	Algebraic Reasoning	Number Sense	Geometry	Measurement	Data Analysis and Probability
Mathematics	--	1.13	1.13	1.11	1.08	1.10
Algebraic Reasoning	0.75	--	0.94	0.92	0.90	0.91
Number Sense	0.85	0.54	--	0.93	0.97	0.92
Geometry	0.79	0.50	0.57	--	0.88	0.89
Measurement	0.77	0.50	0.60	0.52	--	0.86
Data Analysis and Probability	0.80	0.50	0.57	0.53	0.51	--

Table 2.2.Q
Standards Intercorrelation: Grade 8 Mathematics

	Mathematics	Algebraic Reasoning	Number Sense	Geometry	Measurement	Data Analysis and Statistics
Mathematics	--	1.07	1.11	1.11	1.11	1.02
Algebraic Reasoning	0.79	--	0.89	0.85	0.90	0.82
Number Sense	0.79	0.53	--	0.92	0.98	0.87
Geometry	0.73	0.47	0.49	--	0.95	0.84
Measurement	0.88	0.59	0.62	0.55	--	0.87
Data Analysis and Statistics	0.79	0.53	0.54	0.48	0.60	--

Table 2.2.R
Standards Interrelation: Grade 5 Science

	Science	Observe and Measure	Classify	Experiment	Interpret and Communicate
Science	--	1.14	1.14	1.14	1.12
Observe and Measure	0.82	--	1.00	1.00	0.99
Classify	0.83	0.59	--	1.01	0.99
Experiment	0.85	0.61	0.62	--	0.99
Interpret and Communicate	0.90	0.64	0.65	0.68	--

Table 2.2.S
Standards Interrelation: Grade 8 Science

	Science	Observe and Measure	Classify	Experiment	Interpret and Communicate
Science	--	1.17	1.19	1.15	1.13
Observe and Measure	0.79	--	1.03	0.99	1.00
Classify	0.76	0.51	--	0.99	1.00
Experiment	0.88	0.59	0.56	--	0.97
Interpret and Communicate	0.87	0.60	0.57	0.66	--

Table 2.2.T
Standards Interrelation: Grade 5 Social Studies

	Social Studies	Early Exploration	Colonial America	American Revolution	Early Federal Period	Geographic Skills
Social Studies	--	1.15	1.11	1.10	1.09	1.07
Early Exploration	0.74	--	1.03	1.01	0.99	0.95
Colonial America	0.84	0.56	--	0.98	0.97	0.93
American Revolution	0.85	0.56	0.64	--	0.98	0.93
Early Federal Period	0.76	0.50	0.58	0.59	--	0.89
Geographic Skills	0.88	0.56	0.65	0.66	0.58	--

Table 2.2.U
Standards Interrelation: Grade 7 Geography

	Geography	Geographic Tools	Regions	Physical Systems	Human Systems	Human /Environment Interaction	Geography Skills
Geography	--	1.13	1.17	1.18	1.19	1.13	1.13
Geographic Tools	0.65	--	0.97	1.00	0.98	0.96	1.02
Regions	0.84	0.47	--	1.01	1.01	0.96	1.01
Physical Systems	0.75	0.42	0.53	--	0.99	0.98	1.02
Human Systems	0.77	0.42	0.55	0.47	--	0.99	1.01
Human/Environment Interaction	0.80	0.45	0.56	0.51	0.53	--	0.97
Geography Skills	0.74	0.45	0.56	0.50	0.50	0.52	--

Table 2.2.V
Standards Intercorrelation: Grade 8 U. S. History

	History	HA	HB	HC	HD	HE	HF	HG	HH	HI
History	--	1.13	1.07	1.12	1.14	1.14	1.08	1.22	1.13	1.05
HA	0.76	--	0.98	1.02	1.03	1.02	0.97	1.07	0.98	0.93
HB	0.77	0.54	--	0.93	0.96	0.98	0.95	0.96	0.95	0.90
HC	0.73	0.50	0.49	--	1.02	1.03	0.93	1.05	1.00	0.93
HD	0.74	0.50	0.51	0.48	--	0.99	0.97	1.03	0.98	0.91
HE	0.67	0.45	0.46	0.45	0.42	--	1.01	1.07	1.03	0.95
HF	0.72	0.49	0.52	0.46	0.47	0.45	--	1.02	0.98	0.94
HG	0.57	0.38	0.37	0.36	0.35	0.33	0.36	--	1.03	0.96
HH	0.73	0.48	0.50	0.48	0.47	0.44	0.48	0.36	--	0.92
HI	0.74	0.49	0.51	0.48	0.47	0.44	0.50	0.36	0.48	--

HA – Social Studies Process Skills

HB – Causes of American Revolution

HC – Results of American Revolution

HD – Governing Documents/Early Federal Period

HE – Northern/Southern Economic Growth

HF – Jacksonian Era

HG – Cultural Growth and Reform

HH – Westward Movement

HI – Eve of War

2.3 Analysis of the Writing Tests

The administration of the spring 2010 Writing assessment took place on February 24th and March 3rd. Students at Grades 5 and 8 were given one operational writing prompt. The Grade 5 operational prompt was field-test prompt #7 in 2007; the Grade 8 operational prompt was field-test prompt #9 in 2007. The following sections describe the statistical analyses conducted to place the 2010 operational writing prompts on the scale established in 2006.

2.3.1 Prompt Scoring Formula

The writing score is a weighted composite of five analytic scores that focus on specific domains of writing skills. These skills are listed in Table 2.3.A. Each student's response to a prompt is read by two independent raters; the raters' scores for each domain are averaged. The domain scores range from 1 (the lowest score) to 4 (the highest score).

The raw writing score is calculated as a weighted composite of the average of two independent ratings for each of the five analytic traits:

$$\text{Raw Composite Score (RCS)} = 15 * (0.30 * \text{ID} + 0.25 * \text{OUC} + 0.15 * \text{WC} + 0.15 * \text{SP} + 0.15 * \text{GUM})$$

Table 2.3.A
Weights Assigned to Writing Analytic Traits

Writing Analytic Traits	Weight
Ideas and Development (ID)	30%
Organization, Unity, and Coherence (OUC)	25%
Word Choice (WC)	15%
Sentences and Paragraphs (SP)	15%
Grammar, Usage, and Mechanics (GUM)	15%

2.3.2 Statistical Adjustments to Scale the Writing Scores

The baseline for each grade's operational writing scale was 2006. To place the 2010 operational prompts on the 2006 scale, transformation constants were obtained to adjust RCS scores for prompt difficulty and for rater-year effects relative to a target distribution. All calculations were performed on the RCS prior to rounding. For reporting, the scaled composite scores (SCS) were then rounded to the nearest integer between 15 and 60.

Adjustment for Prompt Difficulty and Rater-Year Effects

For each of the 2007 field-test prompts, ETS provided a set of unique transformation constants to adjust for both prompt difficulty and rater-year effects. Based on ETS' report, *OCCT Writing: Scaling the 2007 Field-Test Prompts* (ETS, 2007), the following equation was used to adjust the 2010 raw composite scores ($RCS_{Op10, grade}$):

$$SCS_{Op10, grade} = B_{07, grade} \times RCS_{Op10, grade} + A_{07, grade}.$$

Where $SCS_{Op09, grade}$ represents the scaled composite score after adjusting for the 2007 prompt difficulty and rater-year effects; $A_{07, grade}$ and $B_{07, grade}$ are the additive and multiplicative constants (Grade 5: $A_{07, g5} = -0.647451$ $B_{07, g5} = 1.023409$; Grade 8: $A_{07, g8} = -1.572272$ $B_{07, g8} = 1.043021$).

Adjustment for Rater-Year Effects

In 2010, DRC performed a rater drift study similar to the one conducted by ETS in 2007 to adjust for the rater-year effects. DRC's Performance Assessment Services (PAS) staff randomly pulled 510 student responses from 2007 for each grade's prompt and distributed these into the current administration scoring throughout the entire scoring timeframe. The student responses were pulled by lithocode and were only the valid scored responses (i.e., no condition codes such as off-topic present). 2010 scorers then rescored these papers. The lithcodes randomly pulled by PAS were provided to EIS for generating the data files for Psychometric Services (PS) department.

The 2010 rater-year effect constants, $C_{10, grade}$ and $D_{10, grade}$, were determined by using the means (M) and standard deviations (S) of the 2007 raw composite scores and the 2010 rescored raw composite scores ($RS10$) as calculated below for each grade:

$$C_{10, grade} = M_{07, grade} - M_{RS10, grade} \times D_{10, grade}$$

$$D_{10, grade} = S_{07, grade} / S_{RS10, grade}$$

The formula for the 2010 rater-year effects adjusted score is:

$$SCS_{Op10, grade} = D_{10, grade} \times RCS_{Op10, grade} + C_{10, grade}$$

Once the transformation constants are applied to the 2010 rescored raw composite scores, the mean and standard deviation of the adjusted 2010 scores should be the same as the 2007 mean and standard deviation.

A Compound Adjustment

Following the calculation of the 2010 transformation constants, compound adjustments were made to the 2010 operational raw composite scores. The generic formula for producing the final 2010 scaled composite score ($SCS_{Op10, grade}$) is:

$$SCS_{Op10, grade} = B_{07, grade} \times (D_{10, grade} \times RCS_{Op10, grade} + C_{10, grade}) + A_{07, grade}$$

To simplify the calculation, transformation constants for each of the Grades 5 and 8 were calculated as below:

$$E_{10, grade} = B_{07, grade} \times C_{10, grade} + A_{07, grade}$$

$$F_{10, grade} = D_{10, grade} \times B_{07, grade}$$

The following formula was used to calculate the final scaled composite scores. The calculated values are rounded to the nearest whole integer. Resulting values outside of the 15-60 range are set to the nearest bound.

$$SCS_{Op10, grade} = F_{10, grade} \times RCS_{Op10, grade} + E_{10, grade}$$

The scaled composite score will be converted to the performance level using Table 2.3.B.

Table 2.3.B
Scaled Score Ranges for Each Achievement Level

GRADE 5 SCALED COMPOSITE SCORE	GRADE 8 SCALED COMPOSITE SCORE	Performance Level
54 – 60	54 – 60	Advanced
36 – 53	36 – 53	Satisfactory
26 – 35	25 – 35	Limited Knowledge
15 – 25	15 – 24	Unsatisfactory
Unscorable	Unscorable	Unsatisfactory

Summary statistics for the scaling analysis of the operational writing prompts are provided in Tables 2.3.C to 2.3.E. Table 2.3.C provides the sample means and standard deviations used to calculate the transformation constants for each grade. The results indicate that sampled students in both grades had lower 2010 prompt scores. Because the responses scored were the same across the two years, this indicates that the raters were more strict in 2010.

Table 2.3.C
Sample Means and Standard Deviations Used for Calculating Constants

Grade	Statistic	Raters	
		2007	2010
5	N	508	508
	MIN	15	15
	MAX	60	60
	MEAN*	42.87	41.07
	STD*	9.02	8.41
8	N	509	509
	MIN	15	15
	MAX	60	60
	MEAN*	44.36	41.26
	STD*	7.64	7.29

*Tabled values are rounded for display purposes. Transformations were performed without rounding.

Tables 2.3.D and 2.3.E provide the resulting score distribution statistics with no adjustment, only the ETS adjustment, and the compound DRC and ETS adjustment. Transformation constants are provided at the bottom of the Tables. The 2008 and 2009 score distributions are also provided for comparison. Relative to no adjustment and the ETS only adjustment, the DRC and ETS compound adjustment led to higher mean scores at both grades.

Table 2.3.D
Grade 5 Writing Results

		2010 No Adjustment	2010 ETS Only	2010 DRC & ETS	2009 Scores	2008 Scores
Statistic	N	44994	44994	44994	43665	41988
	MIN	15	15	15	19	17
	MAX	60	60	60	60	60
	MEAN	41.58	41.79	43.67	44.57	44.01
	STD	7.69	7.82	8.25	8.54	8.91
Perf Level %	PL 1 U	2.9	3.1	2.9	3.5	4.6
	PL 2 L	19.8	19.6	15.0	13.7	15.0
	PL 3 S	71.4	71.1	71.5	69.0	67.4
	PL 4 A	6.0	6.3	10.7	13.8	13.0
Constants						
DRC	C	Additive	-1.147132			
	D	Multiplicative	1.071898			
ETS	A	Additive	-0.647451			
	B	Multiplicative	1.023409			
Combined	E	Additive	-1.821436			
	F	Multiplicative	1.096990			

Table 2.3.E
Grade 8 Writing Results

		2010 No Adjustment	2010 ETS Only	2010 DRC & ETS	2009 Scores	2008 Scores
Statistic	N	42153	42153	42153	40962	42271
	MIN	15	15	16	19	18
	MAX	60	60	60	60	60
	MEAN	42.78	42.89	46.19	45.73	45.50
	STD	7.03	7.22	7.39	7.42	7.04
Perf Level %	PL 1 U	2.1	2.3	1.9	2.0	1.8
	PL 2 L	14.4	14.2	9.6	8.8	8.7
	PL 3 S	77.7	77.4	76.4	75.2	78.0
	PL 4 A	5.9	6.1	12.2	14.1	11.6
Constants						
DRC	C	Additive	1.113081			
	D	Multiplicative	1.048035			
ETS	A	Additive	-1.572272			
	B	Multiplicative	1.043021			
Combined	E	Additive	-0.411305			
	F	Multiplicative	1.093122			

2.3.3 Rater Agreement for Operational Writing Prompts

As stated earlier, student responses were rated by two independent raters, and the score for each domain was the average of the two ratings. The average of the two ratings was used for the calculation of the final composite score.

Consistency between the two ratings was evaluated with the following statistics:

- Percentage of exact agreement between raters
- Percentage of adjacent agreement between raters
- Correlation between ratings 1 and 2

Table 2.3.F provides a summary of the rater-agreement analysis for the Grade 5 and Grade 8 operational prompts. Included are the mean and standard deviation of assigned ratings, the percentage of exact and adjacent ratings, and the correlation between ratings. In Grade 5 writing, the exact agreement rate ranged from 69% to 75%, and the sum of exact plus adjacent agreement rates ranged from 99% to 100%. For Grade 8 Writing, the exact agreement rate ranged from 73% to 79%, and the sum of the exact plus adjacent agreement rates was 100%. The correlations between ratings ranged from 0.63 to 0.67 in Grade 5 and 0.60 to 0.66 in Grade 8. In general, the raters were fairly consistent in each domain.

Table 2.3.F
Inter-rater Agreement for Operational Writing Prompts

Grade	Domain ^a	N	Rating 1		Rating 2		Percent Agreement			Corr ^b
			Mean	SD	Mean	SD	Exact	Adjacent	Exact+ Adjacent	
5	ID	44878	2.79	0.61	2.79	0.61	75	25	100	0.64
	OUC	44878	2.66	0.64	2.66	0.64	73	26	99	0.65
	WC	44878	2.80	0.62	2.79	0.62	73	26	100	0.63
	SP	44878	2.78	0.71	2.77	0.70	70	29	99	0.66
	GUM	44878	2.88	0.71	2.87	0.71	69	31	100	0.67
8	ID	42015	2.85	0.58	2.84	0.58	75	25	100	0.63
	OUC	42015	2.83	0.58	2.82	0.58	76	24	100	0.63
	WC	42005	2.93	0.52	2.91	0.52	79	21	100	0.60
	SP	42015	2.84	0.61	2.83	0.61	75	25	100	0.66
	GUM	42015	2.83	0.61	2.83	0.61	73	27	100	0.63

^a ID=Ideas and Development; OUC=Organization, Unity, and Coherence; WC=Word Choice; SP=Sentences and Grammar; GUM=Grammar, Usage, and Mechanics

^b Pearson correlation between first and second ratings

CHAPTER III. STATE RESULTS

In this section, performance on the OCCT is summarized for the participating Oklahoma student population and for demographic subgroups. All reported results are based on valid scores on the 2010 forms in the final student data received by June 24, 2010. These data differ from the analysis data in several ways: corrections were made to student and school information, invalidations and missing data issues were resolved, and all students who took the standard, but not the equivalent and Braille, OCCTs were included. Thus, final counts of examinees by test differ somewhat from samples used for item and test analysis.

As described in Chapter II, prior to the release of student reports, raw scores were converted to a reporting scale metric. Raw scores on the Multiple-Choice tests were converted to scaled scores using the conversion tables provided in Appendix B. For the Writing tests, analytic scores were converted to composite scores using the formulas provided in the previous section. Achievement level scores were assigned as well using the SDE-established OCCT cut scores.

The means and standard deviations of students' raw scores and scaled or composite scores are shown in Table 3.1. Table 3.2 provides the percentage of students in each achievement category in 2008, 2009 and 2010. Tables 3.3 to 3.21 provide test results by demographic subgroups. Tables B.1 through B.19 provide the raw score, scaled score, CSEM, achievement level, and frequency distributions for each OCCT test.

Table 3.1
Means and Standard Deviations of Students' Raw Scores and Scaled Scores

Content Area	Grade	Valid N	Raw Scores				Scaled Scores			
			Mean	SD	Min	Max	Mean	SD	Min	Max
Reading	3	43737	34.4	8.7	1	50	734.6	81.8	400	990
	4	43021	35.6	8.6	2	50	720.1	73.7	400	990
	5	42510	37.2	8.2	0	50	726.6	78.9	400	990
	6	42343	33.2	9.3	1	50	720.9	78.0	400	990
	7	40887	35.1	8.0	3	50	728.1	70.8	400	990
	8	40264	35.8	8.1	0	50	740.0	81.9	400	990
Mathematics	3	44414	33.1	8.3	1	45	736.8	89.4	400	990
	4	43708	32.4	8.0	0	45	734.4	90.0	400	990
	5	42938	30.4	8.4	0	45	729.4	84.2	400	990
	6	42602	30.3	8.6	0	45	719.7	76.1	400	990
	7	41031	29.3	7.7	0	45	723.7	79.1	400	990
	8	40214	30.6	8.2	1	45	728.1	83.8	400	990
Science	5	43457	31.2	7.8	0	45	775.3	72.2	400	990
	8	41109	29.1	7.6	0	45	772.0	68.5	400	990
Social Studies	5	46604	35.9	10.5	0	60	731.6	76.2	400	990
Geography	7	44634	29.0	7.8	2	45	775.3	87.8	400	990
History	8	43885	28.1	8.6	1	45	733.1	81.6	400	990
Writing	5	44870	41.5	7.7	15	60	43.7	8.2	15	60
	8	42005	42.7	7.0	15	60	46.2	7.4	16	60

^a Mean writing composite scores are reported.

Table 3.2 presents the percentage of students scoring in each of the four achievement levels for all students for the current year and the past two years. It shows that for most of the grades and

subject areas, the percentage of students scoring at or above the Satisfactory/Proficient achievement level increased slightly from 2009 to 2010.

Table 3.2
Percentage of Students Performing within Each Achievement Category in 2008 to 2010

Content Area	Grade	Unsatisfactory			Limited Knowledge			Satisfactory/ Proficient			Advanced			Satisfactory or Advanced		
		2008	2009	2010	2008	2009	2010	2008	2009	2010	2008	2009	2010	2008	2009	2010
Reading	3	2	14	13	11	19	18	83	63	67	4	4	2	87	67	69
	4	1	17	18	6	20	17	88	57	62	4	6	3	92	63	65
	5	4	13	13	12	22	22	74	56	56	10	9	9	84	65	65
	6	7	15	16	12	20	21	72	56	57	9	8	7	81	64	64
	7	5	17	18	17	13	15	64	54	53	13	16	14	78	70	67
	8	4	15	14	14	18	16	73	57	57	9	10	13	82	67	70
Mathematics	3	2	12	11	20	22	20	62	45	48	16	21	21	78	66	70
	4	2	13	13	15	20	20	64	50	51	19	17	16	83	67	67
	5	2	13	12	11	22	20	59	35	36	27	30	32	87	65	68
	6	6	18	18	13	19	18	48	31	31	32	32	33	80	63	64
	7	7	22	20	15	15	15	51	35	37	26	28	28	77	63	65
	8	4	21	18	14	18	16	57	36	39	24	25	27	81	61	66
Science	5	3	4	3	12	12	10	58	58	58	27	26	29	84	84	87
	8	4	3	3	9	11	9	75	69	69	12	17	19	87	86	88
Social Studies	5	12	14	12	20	19	18	46	46	47	22	22	24	68	68	71
Geography	7	3	3	2	17	16	16	67	63	64	13	18	18	80	81	82
History	8	9	9	9	25	23	22	56	58	57	10	10	12	66	68	69

Tables 3.3 to 3.21 present the scaled score and achievement level results for each test by population subgroups. Ethnic category membership is based on identifying one ethnicity; those identifying more than one or none are classified as “Other”. Economically disadvantaged is based on participation in Free and Reduced-Price Lunch.

Table 3.3
Subgroup Results: Grade 3 Reading

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		43737	34.4	8.7	0.89	0.89	734.6	81.8	12.7	18.2	66.7	2.4
Gender	Female	21897	33.8	8.8	0.89	0.89	729.5	82.4	14.3	18.9	64.7	2.1
	Male	21797	34.9	8.5	0.88	0.89	739.9	80.8	11.0	17.5	68.8	2.7
Ethnicity	African American	4606	30.8	8.9	0.88	0.88	702.2	81.0	22.1	23.9	53.1	0.9
	American Indian	8325	33.8	8.5	0.88	0.88	728.8	77.8	12.9	19.7	65.6	1.7
	Hispanic	5271	31.1	8.7	0.88	0.88	704.8	78.5	20.0	25.2	54.0	0.8
	Asian	897	36.4	8.4	0.89	0.89	755.4	85.5	8.7	16.2	70.0	5.1
	White	94	32.5	8.3	0.87	0.87	718.8	81.4	16.0	22.3	56.4	5.3
	Pacific Islander	24255	35.9	8.3	0.88	0.88	748.6	80.0	9.3	15.0	72.5	3.2
	Other	289	33.1	9.0	0.89	0.89	725.0	86.7	13.8	23.2	59.9	3.1
IEP	No	42362	34.7	8.5	0.88	0.88	737.4	80.1	11.5	18.1	68.0	2.5
	Yes	1375	24.8	9.1	0.88	0.88	648.9	85.5	48.9	22.0	29.1	0.1
ELL	No	42929	34.5	8.6	0.89	0.89	736.0	81.1	12.1	17.9	67.5	2.5
	Yes	808	25.8	8.8	0.87	0.87	658.4	83.4	40.1	32.1	27.4	0.5
FLS	No	17347	37.3	7.8	0.88	0.88	762.4	78.4	6.3	12.5	76.9	4.3
	Yes	26390	32.4	8.7	0.88	0.88	716.3	78.8	16.9	21.9	60.0	1.2

Table 3.4
Subgroup Results: Grade 4 Reading

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		43021	35.6	8.6	0.89	0.89	720.1	73.7	17.8	17.4	61.5	3.4
Gender	Female	21373	35.2	8.8	0.90	0.90	717.6	75.4	19.2	17.5	59.9	3.4
	Male	21601	35.9	8.3	0.89	0.89	722.7	71.8	16.4	17.3	63.0	3.4
Ethnicity	African American	4569	32.2	9.0	0.89	0.89	692.0	72.7	29.3	20.9	48.5	1.3
	American Indian	8314	35.0	8.4	0.88	0.89	714.3	70.3	18.9	19.0	60.0	2.2
	Hispanic	4938	32.4	9.0	0.89	0.89	693.6	72.3	27.9	21.4	49.4	1.3
	Asian	852	37.4	8.2	0.89	0.89	736.9	75.2	13.3	14.6	66.7	5.5
	Pacific Islander	77	33.8	9.8	0.92	0.92	706.3	81.1	24.7	15.6	55.8	3.9
	White	23856	37.0	8.1	0.88	0.89	732.6	72.0	13.2	15.4	66.9	4.6
	Other	415	34.5	8.5	0.88	0.88	711.8	72.4	20.5	19.8	56.9	2.9
IEP	No	41461	35.9	8.4	0.89	0.89	722.6	72.2	16.6	17.3	62.7	3.5
	Yes	1560	27.3	9.6	0.89	0.89	653.0	79.6	50.0	19.9	29.5	0.6
ELL	No	42417	35.7	8.5	0.89	0.89	721.2	73.2	17.3	17.3	62.0	3.4
	Yes	604	26.0	8.6	0.86	0.86	644.1	70.4	55.5	22.5	21.9	0.2
FLS	No	17443	38.3	7.5	0.88	0.88	744.4	70.6	9.4	12.9	71.9	5.9
	Yes	25578	33.7	8.7	0.89	0.89	703.6	71.1	23.5	20.4	54.3	1.7

Table 3.5
Subgroup Results: Grade 5 Reading

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		42510	37.2	8.2	0.89	0.89	726.6	78.9	13.0	21.7	56.4	8.9
Gender	Female	21223	36.8	8.6	0.90	0.90	723.0	81.4	14.8	21.4	55.3	8.5
	Male	21238	37.6	7.8	0.88	0.88	730.3	76.1	11.2	22.0	57.5	9.3
Ethnicity	African American	4509	33.9	8.7	0.89	0.89	695.7	76.8	22.3	29.6	44.0	4.0
	American Indian	8199	36.4	8.2	0.89	0.89	718.3	76.5	14.3	24.0	55.1	6.5
	Hispanic	4667	34.2	8.6	0.89	0.89	698.0	75.9	21.6	29.0	45.2	4.2
	Asian	809	39.0	8.0	0.90	0.90	747.4	84.8	9.0	17.8	57.7	15.5
	Pacific Islander	87	36.0	8.7	0.90	0.90	716.1	77.4	19.5	25.3	46.0	9.2
	White	23787	38.6	7.6	0.88	0.88	740.4	76.7	9.1	18.2	61.3	11.4
	Other	452	36.1	8.8	0.90	0.90	716.8	81.2	18.4	21.7	53.1	6.9
IEP	No	40801	37.5	7.9	0.89	0.89	729.8	77.1	11.7	21.4	57.6	9.2
	Yes	1709	28.6	9.7	0.90	0.90	650.7	83.7	43.7	29.2	25.6	1.5
ELL	No	42051	37.3	8.1	0.89	0.89	727.6	78.4	12.6	21.6	56.8	9.0
	Yes	459	26.7	8.6	0.87	0.87	635.7	72.0	52.5	31.2	15.7	0.7
FLS	No	17917	39.9	7.0	0.87	0.87	753.2	74.9	6.1	15.4	63.9	14.6
	Yes	24593	35.2	8.5	0.89	0.89	707.2	76.0	18.0	26.3	50.9	4.8

Table 3.6
Subgroup Results: Grade 6 Reading

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		42343	33.2	9.3	0.90	0.90	720.9	78.0	15.5	20.6	57.1	6.8
Gender	Female	21193	32.6	9.6	0.90	0.90	715.5	80.4	18.2	20.8	54.7	6.3
	Male	21111	33.9	8.9	0.89	0.89	726.4	75.0	12.9	20.3	59.4	7.4
Ethnicity	African American	4548	29.3	9.4	0.89	0.89	688.7	77.0	26.8	27.3	43.2	2.7
	American Indian	8028	32.6	9.1	0.89	0.89	715.3	74.8	16.4	22.1	56.2	5.3
	Hispanic	4485	29.7	9.5	0.89	0.89	692.1	77.9	25.9	25.5	45.6	3.1
	Asian	823	36.5	8.7	0.90	0.90	751.5	79.6	7.8	17.4	58.4	16.4
	Pacific Islander	67	29.4	10.9	0.92	0.92	691.2	92.6	31.3	22.4	41.8	4.5
	White	23734	34.8	8.9	0.89	0.89	733.5	75.8	11.4	17.8	62.1	8.6
	Other	658	32.8	8.6	0.87	0.87	716.2	68.2	13.4	23.6	58.7	4.4
IEP	No	40745	33.6	9.1	0.89	0.89	723.9	76.4	14.2	20.3	58.4	7.1
	Yes	1598	23.8	8.8	0.86	0.86	643.9	78.2	50.3	26.0	22.9	0.8
ELL	No	42019	33.3	9.2	0.90	0.90	721.6	77.4	15.2	20.6	57.4	6.9
	Yes	324	21.7	9.0	0.87	0.87	625.0	87.2	62.7	21.3	14.2	1.9
FLS	No	18785	36.2	8.5	0.89	0.89	745.7	74.0	8.6	14.7	65.6	11.2
	Yes	23558	30.8	9.3	0.89	0.89	701.1	75.3	21.1	25.2	50.3	3.4

Table 3.7
Subgroup Results: Grade 7 Reading

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		40887	35.1	8.0	0.87	0.87	728.1	70.8	17.5	15.4	53.3	13.7
Gender	Female	20091	34.4	8.4	0.88	0.88	722.3	73.9	20.6	15.5	51.1	12.8
	Male	20762	35.8	7.5	0.86	0.86	733.8	67.1	14.5	15.3	55.5	14.7
Ethnicity	African American	4122	31.7	8.3	0.87	0.87	698.4	70.6	29.6	19.5	45.2	5.7
	American Indian	7765	34.2	8.0	0.86	0.86	720.0	68.6	20.2	17.1	52.2	10.6
	Hispanic	4237	32.1	8.4	0.87	0.87	701.8	71.8	28.2	19.1	46.0	6.7
	Asian	831	37.3	7.4	0.86	0.86	749.0	68.4	11.9	11.9	54.0	22.1
	Pacific Islander	73	32.3	9.5	0.90	0.90	705.1	84.4	31.5	16.4	38.4	13.7
	White	23311	36.5	7.5	0.86	0.86	740.2	67.9	12.6	13.6	56.6	17.2
	Other	548	34.8	8.1	0.87	0.87	726.4	72.7	19.2	16.6	51.6	12.6
IEP	No	39468	35.4	7.8	0.87	0.87	730.5	69.5	16.3	15.3	54.3	14.2
	Yes	1419	27.3	8.6	0.87	0.87	662.6	74.0	50.5	20.1	27.3	2.0
ELL	No	40584	35.2	7.9	0.87	0.87	728.8	70.2	17.1	15.4	53.6	13.8
	Yes	303	24.6	8.8	0.87	0.87	638.3	80.2	65.3	13.2	19.5	2.0
FLS	No	19085	37.4	7.1	0.85	0.85	748.4	66.3	9.7	12.0	58.2	20.1
	Yes	21802	33.1	8.2	0.87	0.87	710.3	69.7	24.3	18.4	49.1	8.2

Table 3.8
Subgroup Results: Grade 8 Reading

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		40264	35.8	8.1	0.88	0.88	740.0	81.9	14.1	16.4	56.6	12.9
Gender	Female	20212	35.2	8.3	0.88	0.88	733.2	82.4	16.0	17.2	55.6	11.1
	Male	19986	36.5	7.8	0.87	0.87	747.1	80.8	12.1	15.5	57.7	14.7
Ethnicity	African American	4052	32.1	8.5	0.88	0.88	703.6	80.7	26.0	22.9	45.9	5.3
	American Indian	7513	35.3	7.9	0.87	0.87	733.5	77.5	14.6	18.2	57.1	10.0
	Hispanic	4050	32.0	8.9	0.89	0.89	703.3	84.8	27.3	20.0	47.3	5.4
	Asian	828	38.5	8.0	0.89	0.89	770.7	89.2	8.1	10.6	55.6	25.7
	Pacific Islander	82	32.9	9.8	0.91	0.91	709.5	99.1	25.6	12.2	53.7	8.5
	White	23254	37.2	7.5	0.86	0.86	754.1	78.3	9.6	14.2	60.1	16.1
	Other	485	34.8	8.2	0.87	0.88	728.8	79.9	17.5	18.1	54.4	9.9
IEP	No	39341	36.0	8.0	0.88	0.88	741.6	81.1	13.4	16.2	57.2	13.2
	Yes	923	28.5	9.0	0.88	0.88	670.1	85.6	41.6	23.0	33.0	2.4
ELL	No	40008	35.9	8.0	0.88	0.88	740.8	81.4	13.7	16.3	56.9	13.0
	Yes	256	23.0	7.9	0.83	0.83	620.4	78.4	67.2	19.1	12.9	0.8
FLS	No	19640	38.2	7.1	0.86	0.86	764.5	77.3	7.1	12.1	61.8	19.0
	Yes	20624	33.5	8.3	0.87	0.87	716.6	79.4	20.7	20.4	51.7	7.1

Table 3.9
Subgroup Results: Grade 3 Mathematics

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		44414	33.1	8.3	0.90	0.91	736.8	89.4	11.0	19.5	48.2	21.3
Gender	Female	22370	33.6	8.2	0.90	0.91	742.4	89.8	9.9	18.4	48.2	23.4
	Male	22000	32.6	8.3	0.90	0.90	731.2	88.5	12.1	20.6	48.1	19.2
Ethnicity	African American	4667	29.2	9.1	0.91	0.91	696.9	91.9	22.6	25.5	40.3	11.5
	American Indian	8491	32.6	8.2	0.90	0.90	731.5	86.9	11.7	20.7	48.6	19.0
	Hispanic	5365	31.1	8.5	0.90	0.90	715.5	87.9	15.8	24.1	45.9	14.2
	Asian	934	36.1	7.7	0.91	0.91	773.3	94.6	6.7	12.4	45.1	35.8
	Pacific Islander	97	31.0	9.0	0.91	0.91	717.5	95.3	14.4	27.8	40.2	17.5
	White	24567	34.3	7.7	0.89	0.90	749.7	86.2	7.7	17.1	50.2	25.1
	Other	293	31.5	8.4	0.90	0.90	721.4	88.8	14.3	24.6	43.3	17.7
IEP	No	42459	33.4	8.1	0.90	0.90	739.6	88.5	10.3	19.0	48.7	22.1
	Yes	1955	27.2	8.7	0.89	0.89	675.7	86.4	27.1	30.2	36.9	5.8
ELL	No	43175	33.2	8.2	0.90	0.90	738.0	89.0	10.6	19.3	48.4	21.7
	Yes	1239	28.7	9.1	0.90	0.91	693.3	92.2	24.5	26.8	38.5	10.3
FLS	No	17503	35.5	7.3	0.89	0.89	762.7	85.2	5.7	14.4	49.7	30.1
	Yes	26911	31.5	8.5	0.90	0.90	719.9	88.0	14.5	22.8	47.1	15.6

Table 3.10
Subgroup Results: Grade 4 Mathematics

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		43708	32.4	8.0	0.89	0.89	734.4	90.0	13.1	20.0	50.5	16.4
Gender	Female	21892	32.9	8.0	0.89	0.90	740.0	91.4	12.2	18.6	50.8	18.4
	Male	21767	31.9	7.9	0.89	0.89	728.9	88.1	14.0	21.4	50.2	14.4
Ethnicity	African American	4643	28.7	8.3	0.89	0.89	693.5	87.9	24.4	26.2	42.6	6.8
	American Indian	8434	31.8	7.9	0.89	0.89	727.3	87.1	14.2	21.8	50.5	13.5
	Hispanic	5080	30.4	8.2	0.89	0.89	711.3	87.7	18.4	24.1	47.4	10.1
	Asian	879	36.3	7.1	0.89	0.89	784.9	95.2	5.2	11.1	49.4	34.2
	Pacific Islander	82	29.9	9.0	0.91	0.91	707.7	96.8	26.8	18.3	41.5	13.4
	White	24173	33.6	7.5	0.89	0.89	748.0	87.6	9.7	17.6	52.7	20.0
	Other	417	31.8	7.7	0.88	0.88	726.9	84.5	14.4	21.8	53.0	10.8
IEP	No	41539	32.7	7.8	0.89	0.89	737.7	89.0	12.1	19.5	51.3	17.1
	Yes	2169	26.5	8.1	0.87	0.87	670.8	84.2	32.4	29.2	35.3	3.1
ELL	No	42730	32.5	7.9	0.89	0.89	735.7	89.5	12.7	19.8	50.9	16.6
	Yes	978	27.1	8.7	0.89	0.89	678.0	92.2	31.6	26.9	34.6	7.0
FLS	No	17582	34.9	7.1	0.88	0.88	764.1	87.4	6.8	14.6	53.4	25.3
	Yes	26126	30.7	8.0	0.88	0.89	714.4	86.0	17.4	23.6	48.6	10.4

Table 3.11
Subgroup Results: Grade 5 Mathematics

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		42938	30.4	8.4	0.89	0.89	729.4	84.2	12.3	19.5	36.4	31.8
Gender	Female	21565	30.7	8.6	0.90	0.90	732.5	87.0	12.6	18.0	35.4	34.0
	Male	21324	30.1	8.2	0.89	0.89	726.4	81.2	12.0	21.0	37.4	29.6
Ethnicity	African American	4548	27.1	8.5	0.88	0.89	696.7	83.7	21.5	25.4	34.5	18.6
	American Indian	8279	29.2	8.3	0.88	0.88	717.8	80.7	14.5	22.2	37.5	25.7
	Hispanic	4777	28.3	8.5	0.89	0.89	709.4	83.7	17.3	23.2	36.7	22.8
	Asian	827	34.5	8.2	0.91	0.91	775.7	96.2	6.3	10.6	28.7	54.4
	Pacific Islander	90	28.9	9.0	0.90	0.90	713.2	93.2	15.6	23.3	37.8	23.3
	White	23970	31.7	8.0	0.89	0.89	742.4	81.6	8.9	17.0	36.5	37.6
	Other	447	28.6	8.7	0.89	0.90	711.8	87.8	19.2	18.3	37.4	25.1
IEP	No	40804	30.7	8.2	0.89	0.89	732.8	82.6	11.2	19.0	36.7	33.0
	Yes	2134	23.8	8.5	0.88	0.88	664.7	89.5	32.9	28.7	29.0	9.4
ELL	No	42272	30.5	8.3	0.89	0.89	730.5	83.7	11.9	19.4	36.5	32.2
	Yes	666	23.5	8.6	0.88	0.88	661.1	89.4	36.3	27.9	24.9	10.8
FLS	No	18035	33.0	7.7	0.88	0.89	756.1	81.1	6.5	13.9	35.2	44.5
	Yes	24903	28.4	8.3	0.88	0.88	710.0	81.1	16.6	23.6	37.2	22.7

Table 3.12
Subgroup Results: Grade 6 Mathematics

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		42602	30.3	8.6	0.89	0.90	719.7	76.1	18.1	18.2	30.8	32.8
Gender	Female	21428	30.6	8.7	0.90	0.90	722.2	78.6	18.0	17.5	29.6	34.9
	Male	21136	30.0	8.4	0.89	0.89	717.3	73.4	18.2	19.0	32.0	30.7
Ethnicity	African American	4571	26.7	8.8	0.89	0.89	689.3	77.0	31.7	21.1	27.2	19.9
	American Indian	8075	29.4	8.4	0.88	0.89	712.4	72.6	20.0	19.7	32.3	28.0
	Hispanic	4563	28.1	8.8	0.89	0.89	700.7	77.5	25.4	21.1	29.3	24.2
	Asian	843	35.4	7.8	0.90	0.90	769.5	80.8	8.2	9.8	20.3	61.7
	Pacific Islander	71	28.3	9.5	0.91	0.91	706.4	85.1	25.4	28.2	19.7	26.8
	White	23822	31.5	8.2	0.89	0.89	730.2	73.5	13.8	16.9	31.6	37.7
	Other	657	29.3	8.6	0.89	0.89	709.6	77.1	20.4	18.9	33.6	27.1
IEP	No	40656	30.6	8.4	0.89	0.89	722.9	74.6	16.7	18.0	31.3	34.0
	Yes	1946	22.7	8.0	0.86	0.86	654.3	77.7	48.4	23.9	19.8	7.9
ELL	No	42233	30.4	8.6	0.89	0.89	720.3	75.7	17.8	18.2	30.9	33.0
	Yes	369	22.5	9.2	0.89	0.89	652.4	91.7	51.5	21.7	15.2	11.7
FLS	No	18838	32.9	7.9	0.89	0.89	742.1	73.2	10.2	14.4	31.0	44.4
	Yes	23764	28.2	8.5	0.88	0.89	702.0	73.6	24.4	21.3	30.7	23.6

Table 3.13
Subgroup Results: Grade 7 Mathematics

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		41031	29.3	7.7	0.87	0.87	723.7	79.1	20.1	15.2	37.1	27.6
Gender	Female	20276	29.4	7.8	0.87	0.87	725.2	80.5	19.6	14.7	37.4	28.3
	Male	20720	29.1	7.7	0.86	0.87	722.2	77.7	20.6	15.7	36.7	27.0
Ethnicity	African American	4109	26.0	8.0	0.87	0.87	691.4	83.1	35.1	16.8	31.9	16.2
	American Indian	7802	28.2	7.5	0.85	0.86	712.7	76.0	23.0	17.4	37.8	21.7
	Hispanic	4302	26.8	7.8	0.86	0.86	698.8	80.0	30.1	17.2	34.4	18.3
	Asian	845	33.7	7.2	0.87	0.87	772.7	80.9	7.1	10.2	32.8	49.9
	Pacific Islander	77	26.8	8.6	0.89	0.89	699.8	87.9	36.4	14.3	23.4	26.0
	White	23349	30.5	7.4	0.86	0.86	736.0	75.5	15.1	13.9	38.3	32.7
	Other	547	28.8	7.2	0.84	0.84	719.4	70.9	19.2	18.3	41.0	21.6
IEP	No	39347	29.5	7.6	0.86	0.87	726.4	77.7	18.8	15.1	37.6	28.5
	Yes	1684	22.9	7.5	0.84	0.84	659.4	83.1	49.8	18.1	25.1	7.1
ELL	No	40668	29.3	7.7	0.87	0.87	724.3	78.7	19.8	15.2	37.2	27.8
	Yes	363	22.5	8.0	0.86	0.86	653.3	88.9	53.7	16.8	19.8	9.6
FLS	No	19121	31.5	7.2	0.86	0.86	746.5	74.9	12.0	12.2	37.9	37.9
	Yes	21910	27.3	7.6	0.85	0.86	703.8	77.2	27.2	17.8	36.3	18.7

Table 3.14
Subgroup Results: Grade 8 Mathematics

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		40214	30.6	8.2	0.88	0.88	728.1	83.8	18.1	16.0	38.6	27.3
Gender	Female	20258	30.7	8.4	0.89	0.89	729.7	86.5	18.5	15.5	37.2	28.8
	Male	19888	30.5	8.0	0.87	0.88	726.8	80.7	17.6	16.5	39.9	25.9
Ethnicity	African American	4043	27.0	8.3	0.87	0.87	692.7	83.3	30.7	18.5	37.4	13.4
	American Indian	7490	29.6	8.1	0.87	0.88	718.8	80.8	20.9	17.8	38.0	23.4
	Hispanic	4141	28.0	8.4	0.88	0.88	702.1	84.5	27.4	17.9	37.5	17.1
	Asian	850	36.2	7.2	0.89	0.89	793.8	92.3	6.1	6.6	30.2	57.1
	Pacific Islander	85	29.0	9.5	0.91	0.91	711.6	100.6	24.7	16.5	31.8	27.1
	White	23119	31.7	7.8	0.87	0.88	740.0	80.2	13.7	14.9	39.5	31.9
	Other	486	29.3	8.1	0.87	0.88	715.0	81.2	22.0	17.9	39.3	20.8
IEP	No	39077	30.7	8.1	0.88	0.88	730.0	83.0	17.4	15.9	38.8	27.9
	Yes	1137	24.3	8.1	0.86	0.86	665.6	84.6	43.4	20.3	29.4	6.9
ELL	No	39970	30.6	8.2	0.88	0.88	728.6	83.5	17.9	16.0	38.7	27.4
	Yes	244	22.8	8.2	0.86	0.86	651.2	89.6	50.8	20.9	21.7	6.6
FLS	No	19567	32.8	7.6	0.87	0.88	750.7	80.4	11.0	12.8	39.6	36.6
	Yes	20647	28.4	8.2	0.87	0.87	706.8	81.2	24.9	19.0	37.6	18.5

Table 3.15
Subgroup Results: Grade 5 Science

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		43457	31.2	7.8	0.88	0.88	775.3	72.2	2.7	10.4	57.5	29.3
Gender	Female	21831	31.5	8.1	0.89	0.89	778.5	76.1	3.2	10.3	54.2	32.3
	Male	21576	30.9	7.6	0.87	0.87	772.2	67.9	2.3	10.6	60.9	26.3
Ethnicity	African American	4643	26.9	7.9	0.87	0.87	737.5	71.5	6.6	19.8	61.2	12.4
	American Indian	8393	30.5	7.7	0.87	0.87	769.2	69.3	2.6	11.6	60.6	25.2
	Hispanic	4807	28.1	7.7	0.86	0.86	747.9	69.1	4.7	15.8	64.6	15.0
	Asian	829	33.6	7.6	0.89	0.89	798.8	73.0	1.4	6.8	48.5	43.3
	Pacific Islander	89	29.8	8.4	0.89	0.89	761.1	80.6	5.6	13.5	59.6	21.3
	White	24262	32.8	7.4	0.87	0.87	789.4	69.4	1.7	7.2	54.7	36.3
	Other	434	30.6	7.7	0.87	0.87	770.6	66.2	2.3	12.7	58.3	26.7
IEP	No	40951	31.6	7.7	0.88	0.88	778.8	70.8	2.4	9.4	57.6	30.6
	Yes	2506	24.8	7.8	0.85	0.85	719.2	72.6	9.3	26.8	56.2	7.8
ELL	No	42852	31.3	7.8	0.88	0.88	776.3	71.7	2.6	10.1	57.6	29.6
	Yes	605	22.9	7.1	0.82	0.82	703.3	68.3	10.9	33.2	52.4	3.5
FLS	No	18167	33.9	7.0	0.87	0.87	800.0	67.9	1.1	5.4	51.1	42.5
	Yes	25290	29.2	7.8	0.87	0.87	757.6	70.0	3.9	14.1	62.2	19.8

Table 3.16
Subgroup Results: Grade 8 Science

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		41109	29.1	7.6	0.86	0.86	772.0	68.5	3.4	9.1	68.9	18.6
Gender	Female	20765	29.3	7.9	0.87	0.87	774.0	71.8	3.8	9.2	66.0	20.9
	Male	20295	28.9	7.3	0.85	0.85	770.1	64.8	3.0	9.0	71.8	16.3
Ethnicity	African American	4177	25.2	7.3	0.83	0.83	738.3	66.7	7.2	16.9	69.8	6.1
	American Indian	7698	28.2	7.3	0.84	0.84	764.8	64.0	3.4	10.0	72.7	13.9
	Hispanic	4185	25.9	7.7	0.85	0.85	743.4	72.8	7.4	14.6	68.8	9.1
	Asian	855	32.4	7.7	0.88	0.88	801.0	72.4	1.8	6.0	56.4	35.9
	Pacific Islander	84	27.1	7.9	0.86	0.86	752.7	73.8	7.1	13.1	66.7	13.1
	White	23601	30.6	7.3	0.85	0.86	784.6	65.1	2.1	6.6	67.8	23.6
	Other	509	28.6	7.3	0.84	0.85	766.3	70.1	3.3	8.4	74.7	13.6
IEP	No	39236	29.5	7.5	0.86	0.86	775.0	66.8	3.0	8.3	69.3	19.4
	Yes	1873	22.0	6.9	0.81	0.81	709.3	73.0	12.5	26.5	58.5	2.4
ELL	No	40724	29.2	7.6	0.86	0.86	772.9	67.8	3.2	8.9	69.1	18.8
	Yes	385	19.1	6.4	0.76	0.76	680.1	79.1	23.4	33.2	42.1	1.3
FLS	No	19846	31.4	7.1	0.85	0.85	792.0	64.2	1.6	5.3	66.0	27.1
	Yes	21263	27.0	7.5	0.85	0.85	753.4	67.1	5.1	12.7	71.5	10.7

Table 3.17
Subgroup Results: Grade 5 Social Studies

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		46604	35.9	10.5	0.89	0.89	731.6	76.2	11.9	17.6	46.7	23.8
Gender	Female	23825	36.6	10.9	0.90	0.90	735.7	80.3	12.3	15.7	44.5	27.4
	Male	22722	35.2	10.0	0.88	0.88	727.4	71.4	11.5	19.5	48.9	20.1
Ethnicity	African American	5136	30.4	10.1	0.88	0.88	691.7	81.1	24.3	26.6	38.8	10.4
	American Indian	9104	35.1	10.1	0.88	0.88	726.7	72.4	12.1	19.3	48.5	20.1
	Hispanic	5162	32.3	9.9	0.87	0.87	706.8	74.4	17.5	24.3	45.4	12.8
	Asian	845	39.8	10.7	0.91	0.91	758.7	77.0	7.7	11.8	41.2	39.3
	Pacific Islander	93	35.2	10.9	0.90	0.90	724.3	86.2	15.1	15.1	49.5	20.4
	White	25787	37.9	10.2	0.89	0.89	745.4	72.4	8.4	14.0	48.0	29.5
	Other	477	35.5	10.3	0.89	0.89	729.3	73.2	13.6	16.4	47.8	22.2
IEP	No	41255	37.0	10.1	0.89	0.89	740.0	71.9	9.0	15.9	48.7	26.4
	Yes	5349	27.1	8.9	0.84	0.84	667.1	77.6	34.5	30.4	30.9	4.2
ELL	No	45938	36.0	10.5	0.89	0.89	732.6	75.8	11.6	17.3	46.9	24.1
	Yes	666	26.5	8.3	0.82	0.82	664.7	72.4	35.4	33.5	27.5	3.6
FLS	No	18842	40.0	9.8	0.89	0.89	759.7	69.0	5.3	10.7	47.2	36.7
	Yes	27762	33.1	10.0	0.88	0.88	712.5	74.9	16.4	22.2	46.3	15.1

Table 3.18
Subgroup Results: Grade 7 Geography

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		44634	29.0	7.8	0.86	0.86	775.3	87.8	2.4	16.0	63.6	18.0
Gender	Female	22539	29.3	8.0	0.87	0.87	779.2	92.3	2.7	15.8	60.9	20.7
	Male	22025	28.6	7.4	0.85	0.85	771.6	82.6	2.1	16.3	66.3	15.3
Ethnicity	African American	4644	25.0	7.8	0.85	0.85	731.4	89.7	5.6	28.8	58.3	7.2
	American Indian	8659	28.2	7.5	0.84	0.85	766.5	82.9	2.2	17.8	66.2	13.8
	Hispanic	4639	26.3	7.8	0.85	0.85	745.5	88.3	4.3	24.9	61.1	9.8
	Asian	857	32.8	7.0	0.85	0.85	819.1	81.1	0.5	7.0	56.9	35.6
	Pacific Islander	83	26.4	8.8	0.89	0.89	743.6	111.8	8.4	18.1	60.2	13.3
	White	25187	30.3	7.4	0.85	0.85	790.6	84.2	1.6	11.8	64.2	22.4
	Other	565	28.9	7.7	0.86	0.86	773.8	87.8	3.0	14.5	66.5	15.9
IEP	No	41820	29.5	7.5	0.85	0.85	781.6	84.1	1.8	14.0	65.1	19.1
	Yes	2814	20.6	7.1	0.81	0.81	681.8	88.4	12.2	46.4	39.9	1.5
ELL	No	44409	29.0	7.7	0.86	0.86	775.9	87.4	2.3	15.9	63.7	18.1
	Yes	225	19.8	7.3	0.82	0.82	670.5	93.9	14.7	50.2	33.8	1.3
FLS	No	20109	31.7	7.0	0.84	0.84	805.2	80.2	1.0	8.3	63.3	27.4
	Yes	24525	26.8	7.7	0.85	0.85	750.9	86.1	3.6	22.4	63.8	10.3

Table 3.19
Subgroup Results: Grade 8 History

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		43885	28.1	8.6	0.89	0.89	733.1	81.6	9.1	22.1	56.6	12.2
Gender	Female	22538	28.6	8.9	0.90	0.90	737.6	85.6	9.5	20.0	56.0	14.4
	Male	21289	27.5	8.3	0.88	0.88	728.4	76.9	8.7	24.2	57.3	9.8
Ethnicity	African American	4590	24.2	8.3	0.87	0.87	696.6	81.2	16.8	32.6	46.3	4.2
	American Indian	8330	27.1	8.4	0.88	0.88	724.5	79.1	9.9	25.1	55.7	9.3
	Hispanic	4447	25.2	8.4	0.87	0.87	707.1	79.7	14.3	28.6	51.2	5.9
	Asian	867	32.1	8.4	0.90	0.90	771.9	84.0	5.0	11.6	56.5	26.9
	Pacific Islander	89	27.1	8.6	0.89	0.88	723.0	76.9	11.2	24.7	53.9	10.1
	White	25011	29.5	8.3	0.88	0.89	746.2	79.0	6.6	18.4	59.7	15.3
	Other	551	27.4	8.3	0.88	0.88	725.7	80.8	10.7	22.1	57.9	9.3
IEP	No	39762	29.0	8.3	0.88	0.88	741.4	77.2	6.7	20.2	59.8	13.3
	Yes	4123	19.6	7.2	0.82	0.82	653.2	80.1	32.8	40.3	25.7	1.2
ELL	No	43465	28.2	8.6	0.89	0.89	734.0	81.1	8.8	21.9	57.0	12.3
	Yes	420	18.3	6.4	0.77	0.77	640.5	76.3	38.8	40.7	20.2	0.2
FLS	No	20595	30.8	8.1	0.88	0.88	757.6	76.9	4.8	15.1	61.6	18.5
	Yes	23290	25.7	8.4	0.87	0.88	711.4	79.5	12.9	28.3	52.2	6.6

Table 3.20
Subgroup Results: Grade 5 Writing

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores ^a		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		44870	41.5	7.7	-	-	43.7	8.2	1.8	15.2	72.3	10.8
Gender	Female	22871	40.0	7.6	-	-	42.0	8.2	2.6	19.5	70.8	7.1
	Male	21950	43.1	7.4	-	-	45.4	7.9	0.9	10.6	73.8	14.6
Ethnicity	African American	4927	39.4	7.7	-	-	41.3	8.4	3.3	21.6	68.7	6.4
	American Indian	8462	40.8	7.5	-	-	42.9	8.1	1.8	17.2	72.4	8.7
	Hispanic	4677	39.7	7.4	-	-	41.8	8.0	2.5	19.4	72.0	6.0
	Asian	803	44.4	7.4	-	-	46.7	7.7	0.2	8.8	72.5	18.4
	Pacific Islander	97	40.8	7.9	-	-	42.8	8.5	3.1	15.5	70.1	11.3
	White	25140	42.4	7.6	-	-	44.6	8.1	1.4	12.6	73.1	13.0
	Other	764	41.8	7.7	-	-	44.0	8.3	1.3	16.2	69.8	12.7
IEP	No	41006	42.3	7.2	-	-	44.5	7.7	0.8	12.4	75.2	11.6
	Yes	3864	33.4	7.9	-	-	34.8	8.7	12.5	44.1	41.6	1.9
ELL	No	44518	41.6	7.7	-	-	43.7	8.2	1.7	15.0	72.4	10.9
	Yes	352	35.4	7.1	-	-	37.0	7.7	7.7	33.8	58.0	0.6
FLS	No	18630	43.9	7.4	-	-	46.2	7.8	0.7	9.0	73.2	17.1
	Yes	26240	39.9	7.4	-	-	41.9	8.1	2.6	19.5	71.6	6.3

^a Weighted composite scores

Table 3.21
Subgroup Results: Grade 8 Writing

Group	Subgroup	Valid N	Raw Scores		Alpha	Stratified Alpha	Scaled Scores ^a		Percent in Achievement Level			
			Mean	SD			Mean	SD	Unsat	Lim	Sat	Adv
Overall		42005	42.7	7.0	-	-	46.2	7.4	0.8	9.7	77.2	12.3
Gender	Female	21541	41.2	7.1	-	-	44.6	7.6	1.3	13.5	76.9	8.3
	Male	20431	44.3	6.5	-	-	47.9	6.7	0.4	5.7	77.4	16.5
Ethnicity	African American	4326	40.4	7.0	-	-	43.7	7.6	1.3	16.1	76.3	6.3
	American Indian	7300	42.1	7.0	-	-	45.5	7.5	1.1	10.7	78.0	10.1
	Hispanic	3954	41.0	6.8	-	-	44.4	7.3	1.2	13.6	78.6	6.7
	Asian	834	45.7	7.0	-	-	49.2	6.9	0.4	4.7	69.2	25.8
	Pacific Islander	77	42.6	7.1	-	-	46.1	7.5	1.3	11.7	77.9	9.1
	White	24025	43.5	6.9	-	-	47.0	7.2	0.6	7.8	77.0	14.5
	Other	1489	42.8	6.9	-	-	46.2	7.2	0.8	8.4	79.0	11.8
IEP	No	38846	43.4	6.5	-	-	47.0	6.8	0.4	7.1	79.3	13.2
	Yes	3159	34.1	7.2	-	-	36.8	7.9	6.2	41.7	51.2	1.0
ELL	No	41649	42.8	7.0	-	-	46.3	7.3	0.8	9.4	77.4	12.4
	Yes	356	35.1	7.1	-	-	38.0	7.8	4.8	38.2	55.3	1.7
FLS	No	20199	44.5	6.6	-	-	48.1	6.7	0.4	5.6	76.5	17.6
	Yes	21806	41.1	7.0	-	-	44.4	7.5	1.3	13.5	77.8	7.4

^a Weighted composite scores

CHAPTER IV. PERFORMANCE STANDARDS

Performance standards represent the criteria which specify a minimum score a student must achieve on the statewide assessment to be placed into a given performance level. In Oklahoma, four performance levels (i.e., unsatisfactory, limited knowledge, satisfactory, and advanced) were previously established for grades 5 and 8 in 2001, for grades 3 and 4 in 2005, and for grades 6 and 7 in 2006. However, to increase rigor by raising standards for grades 3 through 8 student's achievement on the OCCT as a means to be more competitive at the national and international levels, to vertically align proficiency expectations for students on the OCCT tests across grades 3 through 8, and to align student expectations on the OCCT more closely with student expectations for the National Assessment of Educational Progress (NAEP), revised performance standards (unsatisfactory, limited knowledge, proficient, and advanced) for reading and mathematics were established in 2009.

The workshop to set new academic achievement level cutpoints for grades 3 through 8 in reading and mathematics was held June 15-18, 2009 in Oklahoma City. Thirty seven educational stakeholders from Oklahoma participated in recommending cut scores for the OCCT. Committee members were primarily selected to span grades 3 through 8, although a small number of higher education teachers and those from the business community who are knowledgeable of education in Oklahoma were also selected. The standard setting method known as the Bookmark procedure (Lewis, Mitzel, & Green, 1996), which is the same procedure used in the previous setting of performance level cut scores in reading and mathematics for grades 3 through 8, was employed. The details of the standard setting materials, procedures, methods, and results were reported in the *OCCT Standard Setting: Technical Report* (SDE, 2009). Table 4.1 summarizes the final scaled score ranges for the achievement levels.

Table 4.1 Final Scaled Score Ranges for Reading and Mathematics

Subject	Grade	Unsatisfactory	Limited Knowledge	Proficient	Advanced
Reading	3	400-648	649-699	700-890	891-990
	4	400-657	658-699	700-844	845-990
	5	400-640	641-699	700-829	830-990
	6	400-646	647-699	700-827	828-990
	7	400-667	668-699	700-801	802-990
	8	400-654	655-699	700-832	833-990
Mathematics	3	400-635	636-699	700-797	798-990
	4	400-638	639-699	700-815	816-990
	5	400-641	642-699	700-766	767-990
	6	400-659	660-699	700-753	754-990
	7	400-666	667-699	700-765	766-990
	8	400-661	662-699	700-770	771-990

REFERENCES

- Brennan, R.L. (2004). BB-CLASS [Computer Software]. Iowa City, IA.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Harcourt Brace Jovanovich College Publishers.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297 – 334.
- Data Recognition Corporation (2008). iTEMs [Computer Software]. Maple Grove, MN.
- Dorans, N. J., & Holland, P.W. (1993). DIF Detection and description: Mantel-Haenszel and standardization. In P. W. Holland & H. Wainer (Ed.), *Differential item functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum.
- Hanson, B. A., & Brennan, R. L. (1990). An investigation of classification consistency indexes estimated under alternative strong true score models. *Journal of Educational Measurement*, 27, 345 – 359.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 129–145). Hillsdale, NJ: Lawrence Erlbaum.
- Kolen, M.J., Zeng, L., & Hanson, B.A. (1996). Conditional standard errors of measurement for scale scores using IRT. *Journal of Educational Measurement*, 2, 129-140.
- Lewis, D. M., Mitzel, H. C. & Green, D. R. (1996). Standard Setting: A Bookmark Approach. In D. R. Green (Chair), IRT-based standard-setting procedures utilizing behavioral anchoring. Symposium conducted at the Council of Chief State School Officers National Conference on Large Scale Assessment, Phoenix, AZ.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32(2), 179 – 197.
- Lord, F. M. (1965). A strong true-score theory with applications. *Psychometrika*, 30, 239 – 297.
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22, 719–748.
- Muraki, E., & Bock, R.D. (2003). PARSCALE 4 [Computer Software]. Chicago: Scientific Software International.
- Qualls, A. L. (1995). Estimating the Reliability of a Test Containing Multiple Item Formats. *Applied Measurement in Education*, 8, 111-120.
- SDE (2009). *OCCT Standard Setting: Technical Report*. OK: Oklahoma City.
- Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7(2), 201–210.

Thissen, D., Pommerich, M., Billeaud, K., & Williams, V. S. L. (1995). Item response theory for scores on tests including polytomous items with ordered responses. *Applied Psychological Measurement, 19*, 39-49.

Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAPE history assessment. *Journal of Educational Measurement, 26*, 55-66.

APPENDIX A. DATA REVIEW RESULTS

The items with poor statistics were flagged for SDE's review. The results of the item data review are shown below in Table A.

Table A
Data Review Results

Subject	Grade	Accept	Accept W/R*	Accept Total	Percent Accept	Reject	Total
Mathematics	3						80
	4						80
	5						80
	6						80
	7						80
	8						80
Reading	3						80
	4						80
	5						80
	6						80
	7						80
	8						80
Social Studies	5						80
	7						80
	8						80
Science	5						80
	8						80

* Items may be edited and returned to the pool for future field testing.

APPENDIX B. RAW-TO-SCALED SCORE CONVERSION TABLES AND FREQUENCY DISTRIBUTIONS*

* The frequency analysis was based on the final data files with students who took the standard OCCT.

Table B.1
Raw-to-Scaled Score Table and Frequency Distribution: Grade 3 Reading

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	55	U	0	0.00	0	0.00
1	400	55	U	1	0.00	1	0.00
2	400	55	U	1	0.00	2	0.00
3	400	55	U	1	0.00	3	0.01
4	400	55	U	2	0.00	5	0.01
5	400	55	U	4	0.01	9	0.02
6	400	55	U	9	0.02	18	0.04
7	400	55	U	13	0.03	31	0.07
8	400	55	U	28	0.06	59	0.13
9	400	55	U	57	0.13	116	0.27
10	446	62	U	80	0.18	196	0.45
11	491	66	U	121	0.28	317	0.72
12	519	66	U	154	0.35	471	1.08
13	540	63	U	196	0.45	667	1.53
14	556	58	U	262	0.60	929	2.12
15	569	52	U	296	0.68	1225	2.80
16	581	46	U	320	0.73	1545	3.53
17	592	41	U	394	0.90	1939	4.43
18	602	37	U	461	1.05	2400	5.49
19	611	33	U	482	1.10	2882	6.59
20	619	31	U	549	1.26	3431	7.84
21	628	29	U	649	1.48	4080	9.33
22	636	28	U	688	1.57	4768	10.90
23	643	26	U	767	1.75	5535	12.66
24	651	26	L	872	1.99	6407	14.65
25	658	25	L	1009	2.30	7416	16.95
26	666	25	L	1079	2.47	8495	19.42
27	673	24	L	1106	2.53	9601	21.95
28	680	24	L	1178	2.69	10779	24.64
29	687	24	L	1289	2.95	12068	27.59
30	694	23	L	1420	3.25	13488	30.84
31	702	23	S	1432	3.27	14920	34.11
32	709	23	S	1570	3.59	16490	37.70
33	716	23	S	1665	3.81	18155	41.51
34	724	24	S	1747	3.99	19902	45.50
35	731	24	S	1891	4.32	21793	49.83
36	739	24	S	1826	4.18	23619	54.00
37	747	25	S	1903	4.35	25522	58.35
38	755	25	S	1954	4.47	27476	62.82
39	764	26	S	2018	4.61	29494	67.43
40	774	27	S	2006	4.59	31500	72.02
41	784	29	S	1913	4.37	33413	76.39
42	795	31	S	1877	4.29	35290	80.69
43	807	33	S	1908	4.36	37198	85.05
44	821	37	S	1764	4.03	38962	89.08
45	837	41	S	1512	3.46	40474	92.54
46	856	46	S	1291	2.95	41765	95.49
47	881	50	S	909	2.08	42674	97.57
48	916	49	A	620	1.42	43294	98.99
49	978	36	A	333	0.76	43627	99.75
50	990	36	A	110	0.25	43737	100.00

Table B.2
Raw-to-Scaled Score Table and Frequency Distribution: Grade 4 Reading

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	56	U	0	0.00	0	0.00
1	400	56	U	0	0.00	0	0.00
2	400	56	U	2	0.00	2	0.00
3	400	56	U	1	0.00	3	0.01
4	400	56	U	1	0.00	4	0.01
5	400	56	U	3	0.01	7	0.02
6	400	56	U	7	0.02	14	0.03
7	400	56	U	12	0.03	26	0.06
8	400	56	U	26	0.06	52	0.12
9	400	56	U	49	0.11	101	0.23
10	466	62	U	81	0.19	182	0.42
11	499	64	U	122	0.28	304	0.71
12	522	63	U	123	0.29	427	0.99
13	539	59	U	206	0.48	633	1.47
14	553	53	U	221	0.51	854	1.99
15	565	47	U	238	0.55	1092	2.54
16	576	42	U	272	0.63	1364	3.17
17	585	37	U	357	0.83	1721	4.00
18	594	33	U	368	0.86	2089	4.86
19	602	30	U	386	0.90	2475	5.75
20	610	28	U	451	1.05	2926	6.80
21	617	26	U	482	1.12	3408	7.92
22	624	24	U	570	1.32	3978	9.25
23	631	23	U	579	1.35	4557	10.59
24	637	23	U	662	1.54	5219	12.13
25	644	22	U	733	1.71	5952	13.84
26	650	21	U	810	1.88	6762	15.72
27	656	21	U	897	2.08	7659	17.80
28	662	21	L	1032	2.40	8691	20.20
29	668	21	L	1061	2.47	9752	22.67
30	674	20	L	1125	2.61	10877	25.28
31	681	20	L	1283	2.98	12160	28.27
32	687	20	L	1420	3.30	13580	31.57
33	693	20	L	1546	3.59	15126	35.16
34	700	21	S	1654	3.84	16780	39.01
35	706	21	S	1698	3.95	18478	42.95
36	713	21	S	1876	4.36	20354	47.31
37	721	22	S	1911	4.44	22265	51.75
38	728	22	S	2046	4.76	24311	56.51
39	736	23	S	2142	4.98	26453	61.49
40	744	24	S	2204	5.12	28657	66.61
41	753	25	S	2153	5.00	30810	71.62
42	763	27	S	2204	5.12	33014	76.74
43	774	29	S	2021	4.70	35035	81.44
44	786	32	S	2009	4.67	37044	86.11
45	800	36	S	1774	4.12	38818	90.23
46	817	42	S	1539	3.58	40357	93.81
47	838	48	S	1208	2.81	41565	96.62
48	869	54	A	860	2.00	42425	98.61
49	923	51	A	452	1.05	42877	99.67
50	990	51	A	144	0.33	43021	100.00

Table B.3
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Reading

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	46	U	1	0.00	1	0.00
1	400	46	U	4	0.01	5	0.01
2	400	46	U	6	0.01	11	0.03
3	400	46	U	2	0.00	13	0.03
4	400	46	U	3	0.01	16	0.04
5	400	46	U	5	0.01	21	0.05
6	400	46	U	7	0.02	28	0.07
7	400	46	U	15	0.04	43	0.10
8	400	46	U	13	0.03	56	0.13
9	400	46	U	27	0.06	83	0.20
10	441	53	U	50	0.12	133	0.31
11	476	56	U	70	0.16	203	0.48
12	500	56	U	109	0.26	312	0.73
13	518	54	U	126	0.30	438	1.03
14	533	50	U	166	0.39	604	1.42
15	545	46	U	185	0.44	789	1.86
16	556	41	U	206	0.48	995	2.34
17	566	37	U	228	0.54	1223	2.88
18	575	33	U	284	0.67	1507	3.55
19	583	31	U	299	0.70	1806	4.25
20	591	28	U	314	0.74	2120	4.99
21	599	27	U	338	0.80	2458	5.78
22	606	26	U	416	0.98	2874	6.76
23	613	25	U	380	0.89	3254	7.65
24	620	24	U	457	1.08	3711	8.73
25	627	23	U	527	1.24	4238	9.97
26	634	23	U	596	1.40	4834	11.37
27	640	23	U	696	1.64	5530	13.01
28	647	23	L	733	1.72	6263	14.73
29	654	22	L	873	2.05	7136	16.79
30	661	22	L	1017	2.39	8153	19.18
31	668	22	L	1063	2.50	9216	21.68
32	675	22	L	1250	2.94	10466	24.62
33	682	22	L	1331	3.13	11797	27.75
34	689	23	L	1434	3.37	13231	31.12
35	697	23	L	1539	3.62	14770	34.74
36	704	23	S	1756	4.13	16526	38.88
37	712	24	S	1980	4.66	18506	43.53
38	721	24	S	2009	4.73	20515	48.26
39	729	25	S	2118	4.98	22633	53.24
40	739	26	S	2272	5.34	24905	58.59
41	748	27	S	2336	5.50	27241	64.08
42	759	28	S	2415	5.68	29656	69.76
43	770	30	S	2420	5.69	32076	75.46
44	783	33	S	2465	5.80	34541	81.25
45	798	37	S	2198	5.17	36739	86.42
46	815	42	S	1988	4.68	38727	91.10
47	837	49	A	1696	3.99	40423	95.09
48	868	54	A	1190	2.80	41613	97.89
49	919	53	A	668	1.57	42281	99.46
50	990	53	A	229	0.54	42510	100.00

Table B.4
Raw-to-Scaled Score Table and Frequency Distribution: Grade 6 Reading

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	58	U	0	0.00	0	0.00
1	400	58	U	3	0.01	3	0.01
2	400	58	U	4	0.01	7	0.02
3	400	58	U	3	0.01	10	0.02
4	400	58	U	6	0.01	16	0.04
5	400	58	U	3	0.01	19	0.04
6	400	58	U	8	0.02	27	0.06
7	400	58	U	23	0.05	50	0.12
8	400	58	U	45	0.11	95	0.22
9	400	58	U	60	0.14	155	0.37
10	461	65	U	108	0.26	263	0.62
11	499	68	U	149	0.35	412	0.97
12	525	67	U	227	0.54	639	1.51
13	546	64	U	315	0.74	954	2.25
14	562	59	U	360	0.85	1314	3.10
15	577	53	U	393	0.93	1707	4.03
16	589	47	U	540	1.28	2247	5.31
17	600	42	U	553	1.31	2800	6.61
18	610	37	U	622	1.47	3422	8.08
19	620	34	U	717	1.69	4139	9.77
20	628	31	U	741	1.75	4880	11.52
21	636	29	U	828	1.96	5708	13.48
22	644	27	U	874	2.06	6582	15.54
23	651	25	L	926	2.19	7508	17.73
24	658	24	L	937	2.21	8445	19.94
25	664	23	L	978	2.31	9423	22.25
26	671	22	L	1080	2.55	10503	24.80
27	677	22	L	1106	2.61	11609	27.42
28	683	21	L	1171	2.77	12780	30.18
29	689	21	L	1235	2.92	14015	33.10
30	695	21	L	1272	3.00	15287	36.10
31	701	21	S	1351	3.19	16638	39.29
32	707	21	S	1403	3.31	18041	42.61
33	714	21	S	1441	3.40	19482	46.01
34	720	21	S	1506	3.56	20988	49.57
35	727	21	S	1677	3.96	22665	53.53
36	733	22	S	1714	4.05	24379	57.58
37	740	22	S	1666	3.93	26045	61.51
38	748	23	S	1740	4.11	27785	65.62
39	756	24	S	1807	4.27	29592	69.89
40	764	25	S	1788	4.22	31380	74.11
41	773	26	S	1762	4.16	33142	78.27
42	783	28	S	1724	4.07	34866	82.34
43	793	30	S	1630	3.85	36496	86.19
44	806	33	S	1595	3.77	38091	89.96
45	820	37	S	1354	3.20	39445	93.16
46	837	42	A	1090	2.57	40535	95.73
47	858	48	A	855	2.02	41390	97.75
48	889	51	A	571	1.35	41961	99.10
49	944	44	A	295	0.70	42256	99.79
50	990	44	A	87	0.21	42343	100.00

Table B.5
Raw-to-Scaled Score Table and Frequency Distribution: Grade 7 Reading

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	49	U	0	0.00	0	0.00
1	400	49	U	0	0.00	0	0.00
2	400	49	U	0	0.00	0	0.00
3	400	49	U	2	0.00	2	0.00
4	400	49	U	2	0.00	4	0.01
5	400	49	U	2	0.00	6	0.01
6	400	49	U	5	0.01	11	0.03
7	400	49	U	9	0.02	20	0.05
8	400	49	U	18	0.04	38	0.09
9	400	49	U	30	0.07	68	0.17
10	439	57	U	41	0.10	109	0.27
11	479	61	U	67	0.16	176	0.43
12	506	62	U	93	0.23	269	0.66
13	527	60	U	123	0.30	392	0.96
14	544	56	U	145	0.36	537	1.32
15	559	51	U	176	0.43	713	1.75
16	572	46	U	241	0.59	954	2.34
17	583	41	U	311	0.76	1265	3.10
18	593	37	U	327	0.80	1592	3.90
19	603	34	U	386	0.94	1978	4.84
20	612	31	U	405	0.99	2383	5.83
21	620	29	U	481	1.18	2864	7.01
22	628	28	U	570	1.39	3434	8.40
23	636	27	U	621	1.52	4055	9.92
24	643	26	U	613	1.50	4668	11.42
25	651	25	U	690	1.69	5358	13.11
26	658	24	U	856	2.09	6214	15.20
27	665	24	U	938	2.29	7152	17.49
28	672	23	L	1076	2.63	8228	20.13
29	679	23	L	1137	2.78	9365	22.91
30	685	23	L	1258	3.08	10623	25.98
31	692	22	L	1366	3.34	11989	29.32
32	699	22	L	1471	3.60	13460	32.92
33	706	23	S	1597	3.91	15057	36.83
34	713	23	S	1701	4.16	16758	40.99
35	721	23	S	1921	4.70	18679	45.69
36	728	23	S	2015	4.93	20694	50.61
37	736	24	S	2025	4.95	22719	55.57
38	744	24	S	2169	5.30	24888	60.87
39	753	25	S	2167	5.30	27055	66.17
40	762	26	S	2146	5.25	29201	71.42
41	771	27	S	2176	5.32	31377	76.74
42	781	28	S	2036	4.98	33413	81.72
43	793	30	S	1856	4.54	35269	86.26
44	805	33	A	1706	4.17	36975	90.43
45	819	37	A	1344	3.29	38319	93.72
46	837	43	A	1064	2.60	39383	96.32
47	858	48	A	797	1.95	40180	98.27
48	889	50	A	431	1.05	40611	99.32
49	944	44	A	219	0.54	40830	99.86
50	990	44	A	57	0.14	40887	100.00

Table B.6
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Reading

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	53	U	1	0.00	1	0.00
1	400	53	U	1	0.00	2	0.00
2	400	53	U	1	0.00	3	0.01
3	400	53	U	2	0.00	5	0.01
4	400	53	U	2	0.00	7	0.02
5	400	53	U	1	0.00	8	0.02
6	400	53	U	2	0.00	10	0.02
7	400	53	U	3	0.01	13	0.03
8	400	53	U	15	0.04	28	0.07
9	400	53	U	21	0.05	49	0.12
10	400	53	U	38	0.09	87	0.22
11	453	59	U	53	0.13	140	0.35
12	488	63	U	67	0.17	207	0.51
13	513	62	U	116	0.29	323	0.80
14	532	59	U	160	0.39	483	1.20
15	548	55	U	180	0.45	663	1.64
16	561	50	U	215	0.53	878	2.18
17	573	45	U	240	0.60	1118	2.77
18	584	41	U	315	0.78	1433	3.56
19	594	37	U	353	0.88	1786	4.43
20	604	34	U	393	0.98	2179	5.41
21	613	32	U	429	1.07	2608	6.47
22	622	30	U	478	1.19	3086	7.66
23	630	29	U	535	1.33	3621	8.99
24	638	28	U	671	1.67	4292	10.66
25	646	27	U	683	1.70	4975	12.35
26	654	26	U	697	1.73	5672	14.08
27	661	26	L	845	2.10	6517	16.18
28	669	25	L	939	2.33	7456	18.52
29	676	25	L	1033	2.57	8489	21.08
30	684	25	L	1193	2.96	9682	24.04
31	691	25	L	1227	3.05	10909	27.09
32	699	25	L	1349	3.35	12258	30.44
33	707	25	S	1392	3.46	13650	33.90
34	714	25	S	1514	3.76	15164	37.66
35	722	25	S	1821	4.52	16985	42.18
36	731	26	S	1828	4.54	18813	46.72
37	739	26	S	1838	4.56	20651	51.29
38	748	27	S	2047	5.08	22698	56.37
39	758	28	S	2131	5.29	24829	61.66
40	768	29	S	2182	5.42	27011	67.08
41	779	31	S	2155	5.35	29166	72.44
42	791	32	S	2034	5.05	31200	77.49
43	804	35	S	1986	4.93	33186	82.42
44	819	38	S	1878	4.66	35064	87.08
45	836	42	A	1595	3.96	36659	91.05
46	856	47	A	1495	3.71	38154	94.76
47	882	51	A	1076	2.67	39230	97.43
48	918	50	A	628	1.56	39858	98.99
49	982	36	A	329	0.82	40187	99.81
50	990	36	A	77	0.19	40264	100.00

Table B.7
Raw-to-Scaled Score Table and Frequency Distribution: Grade 3 Mathematics

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	45	U	0	0.00	0	0.00
1	400	45	U	1	0.00	1	0.00
2	400	45	U	1	0.00	2	0.00
3	400	45	U	1	0.00	3	0.01
4	400	45	U	2	0.00	5	0.01
5	400	45	U	2	0.00	7	0.02
6	400	45	U	23	0.05	30	0.07
7	400	45	U	28	0.06	58	0.13
8	400	45	U	45	0.10	103	0.23
9	441	53	U	69	0.16	172	0.39
10	475	57	U	109	0.25	281	0.63
11	500	57	U	151	0.34	432	0.97
12	520	55	U	178	0.40	610	1.37
13	537	52	U	250	0.56	860	1.94
14	552	48	U	328	0.74	1188	2.67
15	565	43	U	318	0.72	1506	3.39
16	577	40	U	415	0.93	1921	4.33
17	589	37	U	464	1.04	2385	5.37
18	599	34	U	480	1.08	2865	6.45
19	609	32	U	614	1.38	3479	7.83
20	619	30	U	663	1.49	4142	9.33
21	628	29	U	752	1.69	4894	11.02
22	637	28	L	774	1.74	5668	12.76
23	646	27	L	876	1.97	6544	14.73
24	654	26	L	1009	2.27	7553	17.01
25	662	25	L	1025	2.31	8578	19.31
26	670	24	L	1101	2.48	9679	21.79
27	678	24	L	1222	2.75	10901	24.54
28	686	23	L	1307	2.94	12208	27.49
29	693	23	L	1340	3.02	13548	30.50
30	701	23	S	1421	3.20	14969	33.70
31	708	23	S	1604	3.61	16573	37.31
32	716	22	S	1644	3.70	18217	41.02
33	724	23	S	1785	4.02	20002	45.04
34	732	23	S	1817	4.09	21819	49.13
35	740	23	S	1912	4.30	23731	53.43
36	749	24	S	2035	4.58	25766	58.01
37	758	25	S	2142	4.82	27908	62.84
38	768	27	S	2266	5.10	30174	67.94
39	779	30	S	2340	5.27	32514	73.21
40	792	34	S	2423	5.46	34937	78.66
41	808	40	A	2502	5.63	37439	84.30
42	827	47	A	2318	5.22	39757	89.51
43	855	55	A	2157	4.86	41914	94.37
44	902	58	A	1621	3.65	43535	98.02
45	990	58	A	879	1.98	44414	100.00

Table B.8
Raw-to-Scaled Score Table and Frequency Distribution: Grade 4 Mathematics

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	49	U	2	0.00	2	0.00
1	400	49	U	1	0.00	3	0.01
2	400	49	U	1	0.00	4	0.01
3	400	49	U	1	0.00	5	0.01
4	400	49	U	5	0.01	10	0.02
5	400	49	U	4	0.01	14	0.03
6	400	49	U	8	0.02	22	0.05
7	400	49	U	11	0.03	33	0.08
8	400	49	U	26	0.06	59	0.13
9	400	49	U	47	0.11	106	0.24
10	442	57	U	76	0.17	182	0.42
11	481	61	U	121	0.28	303	0.69
12	507	61	U	195	0.45	498	1.14
13	527	59	U	209	0.48	707	1.62
14	544	55	U	317	0.73	1024	2.34
15	558	50	U	359	0.82	1383	3.16
16	571	45	U	436	1.00	1819	4.16
17	583	41	U	458	1.05	2277	5.21
18	594	38	U	504	1.15	2781	6.36
19	605	35	U	590	1.35	3371	7.71
20	615	33	U	728	1.67	4099	9.38
21	625	31	U	774	1.77	4873	11.15
22	634	30	U	855	1.96	5728	13.11
23	643	29	L	971	2.22	6699	15.33
24	652	28	L	1077	2.46	7776	17.79
25	661	27	L	1118	2.56	8894	20.35
26	669	26	L	1213	2.78	10107	23.12
27	678	26	L	1333	3.05	11440	26.17
28	686	26	L	1459	3.34	12899	29.51
29	695	25	L	1563	3.58	14462	33.09
30	703	25	S	1707	3.91	16169	36.99
31	712	25	S	1731	3.96	17900	40.95
32	720	25	S	1808	4.14	19708	45.09
33	729	25	S	1896	4.34	21604	49.43
34	738	26	S	1984	4.54	23588	53.97
35	748	26	S	2096	4.80	25684	58.76
36	758	27	S	2113	4.83	27797	63.60
37	769	29	S	2224	5.09	30021	68.69
38	780	30	S	2141	4.90	32162	73.58
39	793	33	S	2164	4.95	34326	78.53
40	808	36	S	2220	5.08	36546	83.61
41	825	41	A	2084	4.77	38630	88.38
42	847	46	A	1882	4.31	40512	92.69
43	876	51	A	1550	3.55	42062	96.23
44	925	49	A	1098	2.51	43160	98.75
45	990	49	A	548	1.25	43708	100.00

Table B.9
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Mathematics

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	51	U	1	0.00	1	0.00
1	400	51	U	1	0.00	2	0.00
2	400	51	U	3	0.01	5	0.01
3	400	51	U	2	0.00	7	0.02
4	400	51	U	8	0.02	15	0.03
5	400	51	U	12	0.03	27	0.06
6	400	51	U	19	0.04	46	0.11
7	400	51	U	43	0.10	89	0.21
8	400	51	U	58	0.14	147	0.34
9	417	62	U	98	0.23	245	0.57
10	477	68	U	177	0.41	422	0.98
11	512	70	U	206	0.48	628	1.46
12	538	68	U	295	0.69	923	2.15
13	558	63	U	358	0.83	1281	2.98
14	575	57	U	442	1.03	1723	4.01
15	590	51	U	541	1.26	2264	5.27
16	602	45	U	643	1.50	2907	6.77
17	614	40	U	714	1.66	3621	8.43
18	625	36	U	793	1.85	4414	10.28
19	635	33	U	877	2.04	5291	12.32
20	644	30	L	909	2.12	6200	14.44
21	653	28	L	1032	2.40	7232	16.84
22	661	27	L	1104	2.57	8336	19.41
23	669	26	L	1195	2.78	9531	22.20
24	677	25	L	1320	3.07	10851	25.27
25	685	24	L	1344	3.13	12195	28.40
26	692	24	L	1468	3.42	13663	31.82
27	700	23	S	1528	3.56	15191	35.38
28	707	23	S	1525	3.55	16716	38.93
29	714	23	S	1648	3.84	18364	42.77
30	722	22	S	1742	4.06	20106	46.83
31	729	22	S	1700	3.96	21806	50.78
32	737	22	S	1871	4.36	23677	55.14
33	744	23	S	1809	4.21	25486	59.36
34	752	23	S	1852	4.31	27338	63.67
35	761	24	S	1935	4.51	29273	68.18
36	769	25	A	1886	4.39	31159	72.57
37	779	26	A	1761	4.10	32920	76.67
38	789	28	A	1745	4.06	34665	80.73
39	801	31	A	1747	4.07	36412	84.80
40	814	35	A	1684	3.92	38096	88.72
41	830	40	A	1489	3.47	39585	92.19
42	850	46	A	1282	2.99	40867	95.18
43	879	51	A	1029	2.40	41896	97.57
44	931	48	A	701	1.63	42597	99.21
45	990	48	A	341	0.79	42938	100.00

Table B.10
Raw-to-Scaled Score Table and Frequency Distribution: Grade 6 Mathematics

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	67	U	1	0.00	1	0.00
1	400	67	U	3	0.01	4	0.01
2	400	67	U	0	0.00	4	0.01
3	400	67	U	2	0.00	6	0.01
4	400	67	U	2	0.00	8	0.02
5	400	67	U	3	0.01	11	0.03
6	400	67	U	17	0.04	28	0.07
7	400	67	U	37	0.09	65	0.15
8	400	67	U	58	0.14	123	0.29
9	400	67	U	90	0.21	213	0.50
10	477	74	U	151	0.35	364	0.85
11	521	76	U	238	0.56	602	1.41
12	549	74	U	320	0.75	922	2.16
13	570	68	U	423	0.99	1345	3.16
14	586	60	U	525	1.23	1870	4.39
15	599	52	U	554	1.30	2424	5.69
16	611	45	U	641	1.50	3065	7.19
17	621	39	U	802	1.88	3867	9.08
18	631	34	U	879	2.06	4746	11.14
19	639	30	U	917	2.15	5663	13.29
20	647	28	U	967	2.27	6630	15.56
21	655	26	U	1097	2.57	7727	18.14
22	662	24	L	1212	2.84	8939	20.98
23	669	23	L	1167	2.74	10106	23.72
24	676	22	L	1216	2.85	11322	26.58
25	682	21	L	1329	3.12	12651	29.70
26	689	21	L	1415	3.32	14066	33.02
27	695	20	L	1430	3.36	15496	36.37
28	701	20	S	1529	3.59	17025	39.96
29	707	20	S	1565	3.67	18590	43.64
30	714	20	S	1625	3.81	20215	47.45
31	720	20	S	1578	3.70	21793	51.15
32	726	20	S	1607	3.77	23400	54.93
33	733	20	S	1768	4.15	25168	59.08
34	740	20	S	1667	3.91	26835	62.99
35	747	21	S	1780	4.18	28615	67.17
36	755	22	A	1811	4.25	30426	71.42
37	763	23	A	1820	4.27	32246	75.69
38	772	25	A	1751	4.11	33997	79.80
39	781	27	A	1741	4.09	35738	83.89
40	793	31	A	1662	3.90	37400	87.79
41	806	37	A	1601	3.76	39001	91.55
42	823	45	A	1389	3.26	40390	94.81
43	847	56	A	1127	2.65	41517	97.45
44	887	62	A	765	1.80	42282	99.25
45	990	62	A	320	0.75	42602	100.00

Table B.11
Raw-to-Scaled Score Table and Frequency Distribution: Grade 7 Mathematics

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	61	U	2	0.00	2	0.00
1	400	61	U	2	0.00	4	0.01
2	400	61	U	2	0.00	6	0.01
3	400	61	U	1	0.00	7	0.02
4	400	61	U	1	0.00	8	0.02
5	400	61	U	8	0.02	16	0.04
6	400	61	U	12	0.03	28	0.07
7	400	61	U	27	0.07	55	0.13
8	400	61	U	60	0.15	115	0.28
9	407	61	U	88	0.21	203	0.49
10	475	68	U	129	0.31	332	0.81
11	511	70	U	189	0.46	521	1.27
12	536	68	U	260	0.63	781	1.90
13	555	64	U	320	0.78	1101	2.68
14	571	58	U	403	0.98	1504	3.67
15	586	52	U	489	1.19	1993	4.86
16	598	46	U	567	1.38	2560	6.24
17	610	41	U	701	1.71	3261	7.95
18	621	37	U	739	1.80	4000	9.75
19	632	34	U	896	2.18	4896	11.93
20	642	32	U	988	2.41	5884	14.34
21	651	31	U	1127	2.75	7011	17.09
22	660	29	U	1237	3.01	8248	20.10
23	669	28	L	1414	3.45	9662	23.55
24	678	28	L	1431	3.49	11093	27.04
25	687	27	L	1603	3.91	12696	30.94
26	696	26	L	1789	4.36	14485	35.30
27	704	26	S	1865	4.55	16350	39.85
28	712	25	S	1842	4.49	18192	44.34
29	721	25	S	1942	4.73	20134	49.07
30	729	24	S	1991	4.85	22125	53.92
31	737	24	S	1988	4.85	24113	58.77
32	746	24	S	1886	4.60	25999	63.36
33	754	24	S	1920	4.68	27919	68.04
34	763	24	S	1770	4.31	29689	72.36
35	772	25	A	1708	4.16	31397	76.52
36	781	26	A	1629	3.97	33026	80.49
37	791	27	A	1505	3.67	34531	84.16
38	802	28	A	1411	3.44	35942	87.60
39	813	30	A	1307	3.19	37249	90.78
40	826	34	A	1074	2.62	38323	93.40
41	842	38	A	939	2.29	39262	95.69
42	861	43	A	728	1.77	39990	97.46
43	888	47	A	539	1.31	40529	98.78
44	934	44	A	363	0.88	40892	99.66
45	990	44	A	139	0.34	41031	100.00

Table B.12
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Mathematics

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	68	U	0	0.00	0	0.00
1	400	68	U	2	0.00	2	0.00
2	400	68	U	2	0.00	4	0.01
3	400	68	U	2	0.00	6	0.01
4	400	68	U	2	0.00	8	0.02
5	400	68	U	6	0.01	14	0.03
6	400	68	U	7	0.02	21	0.05
7	400	68	U	25	0.06	46	0.11
8	400	68	U	44	0.11	90	0.22
9	400	68	U	76	0.19	166	0.41
10	442	68	U	142	0.35	308	0.77
11	497	72	U	185	0.46	493	1.23
12	530	72	U	221	0.55	714	1.78
13	553	69	U	317	0.79	1031	2.56
14	571	63	U	389	0.97	1420	3.53
15	586	56	U	446	1.11	1866	4.64
16	599	50	U	540	1.34	2406	5.98
17	611	44	U	587	1.46	2993	7.44
18	622	39	U	681	1.69	3674	9.14
19	631	35	U	734	1.83	4408	10.96
20	640	32	U	885	2.20	5293	13.16
21	649	30	U	947	2.35	6240	15.52
22	657	28	U	1052	2.62	7292	18.13
23	665	27	L	1106	2.75	8398	20.88
24	673	26	L	1229	3.06	9627	23.94
25	680	25	L	1344	3.34	10971	27.28
26	688	24	L	1339	3.33	12310	30.61
27	695	24	L	1410	3.51	13720	34.12
28	702	24	S	1589	3.95	15309	38.07
29	710	24	S	1592	3.96	16901	42.03
30	717	23	S	1625	4.04	18526	46.07
31	725	24	S	1742	4.33	20268	50.40
32	733	24	S	1747	4.34	22015	54.74
33	741	24	S	1818	4.52	23833	59.27
34	749	24	S	1909	4.75	25742	64.01
35	758	25	S	1731	4.30	27473	68.32
36	767	26	S	1755	4.36	29228	72.68
37	777	28	A	1780	4.43	31008	77.11
38	788	30	A	1671	4.16	32679	81.26
39	800	32	A	1619	4.03	34298	85.29
40	814	36	A	1429	3.55	35727	88.84
41	830	40	A	1354	3.37	37081	92.21
42	851	46	A	1175	2.92	38256	95.13
43	879	50	A	912	2.27	39168	97.40
44	927	48	A	678	1.69	39846	99.08
45	990	48	A	368	0.92	40214	100.00

Table B.13
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Science

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	86	U	1	0.00	1	0.00
1	400	86	U	5	0.01	6	0.01
2	400	86	U	1	0.00	7	0.02
3	400	86	U	0	0.00	7	0.02
4	400	86	U	4	0.01	11	0.03
5	400	86	U	7	0.02	18	0.04
6	400	86	U	16	0.04	34	0.08
7	400	86	U	21	0.05	55	0.13
8	400	86	U	41	0.09	96	0.22
9	498	86	U	69	0.16	165	0.38
10	547	88	U	114	0.26	279	0.64
11	576	84	U	136	0.31	415	0.95
12	597	76	U	197	0.45	612	1.41
13	614	66	U	258	0.59	870	2.00
14	628	56	U	325	0.75	1195	2.75
15	640	48	L	359	0.83	1554	3.58
16	651	41	L	471	1.08	2025	4.66
17	661	36	L	547	1.26	2572	5.92
18	671	32	L	660	1.52	3232	7.44
19	680	29	L	759	1.75	3991	9.18
20	688	27	L	799	1.84	4790	11.02
21	696	26	L	938	2.16	5728	13.18
22	704	25	S	1056	2.43	6784	15.61
23	711	24	S	1055	2.43	7839	18.04
24	718	23	S	1195	2.75	9034	20.79
25	726	23	S	1285	2.96	10319	23.75
26	733	22	S	1415	3.26	11734	27.00
27	740	22	S	1451	3.34	13185	30.34
28	747	22	S	1603	3.69	14788	34.03
29	754	21	S	1741	4.01	16529	38.04
30	761	21	S	1854	4.27	18383	42.30
31	768	21	S	2003	4.61	20386	46.91
32	776	22	S	2005	4.61	22391	51.52
33	783	22	S	2002	4.61	24393	56.13
34	791	22	S	2060	4.74	26453	60.87
35	800	23	S	2188	5.03	28641	65.91
36	808	23	S	2096	4.82	30737	70.73
37	817	25	A	2110	4.86	32847	75.59
38	827	26	A	2093	4.82	34940	80.40
39	838	28	A	1985	4.57	36925	84.97
40	851	31	A	1795	4.13	38720	89.10
41	865	35	A	1571	3.62	40291	92.71
42	883	38	A	1360	3.13	41651	95.84
43	908	41	A	1004	2.31	42655	98.15
44	950	36	A	575	1.32	43230	99.48
45	990	36	A	227	0.52	43457	100.00

Table B.14
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Science

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	93	U	3	0.01	3	0.01
1	400	93	U	7	0.02	10	0.02
2	400	93	U	6	0.01	16	0.04
3	400	93	U	3	0.01	19	0.05
4	400	93	U	6	0.01	25	0.06
5	400	93	U	8	0.02	33	0.08
6	400	93	U	16	0.04	49	0.12
7	400	93	U	15	0.04	64	0.16
8	400	93	U	42	0.10	106	0.26
9	421	93	U	67	0.16	173	0.42
10	544	99	U	108	0.26	281	0.68
11	585	98	U	139	0.34	420	1.02
12	610	90	U	238	0.58	658	1.60
13	628	79	U	338	0.82	996	2.42
14	642	67	U	407	0.99	1403	3.41
15	654	56	L	491	1.19	1894	4.61
16	665	46	L	605	1.47	2499	6.08
17	675	39	L	711	1.73	3210	7.81
18	685	34	L	919	2.24	4129	10.04
19	694	30	L	1014	2.47	5143	12.51
20	702	28	S	1017	2.47	6160	14.98
21	711	26	S	1198	2.91	7358	17.90
22	719	25	S	1291	3.14	8649	21.04
23	726	25	S	1324	3.22	9973	24.26
24	734	24	S	1535	3.73	11508	27.99
25	742	23	S	1570	3.82	13078	31.81
26	749	23	S	1694	4.12	14772	35.93
27	756	22	S	1670	4.06	16442	40.00
28	764	22	S	1802	4.38	18244	44.38
29	771	22	S	1914	4.66	20158	49.04
30	778	21	S	1892	4.60	22050	53.64
31	785	21	S	1990	4.84	24040	58.48
32	792	21	S	1936	4.71	25976	63.19
33	800	22	S	1991	4.84	27967	68.03
34	807	22	S	1843	4.48	29810	72.51
35	815	22	S	1869	4.55	31679	77.06
36	824	23	S	1769	4.30	33448	81.36
37	833	25	A	1645	4.00	35093	85.37
38	842	26	A	1469	3.57	36562	88.94
39	853	29	A	1325	3.22	37887	92.16
40	866	32	A	1078	2.62	38965	94.78
41	881	35	A	871	2.12	39836	96.90
42	900	38	A	640	1.56	40476	98.46
43	925	38	A	370	0.90	40846	99.36
44	969	30	A	216	0.53	41062	99.89
45	990	30	A	47	0.11	41109	100.00

Table B.15
Raw-to-Scaled Score Table and Frequency Distribution: Grade 5 Social Studies

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	70	U	1	0.00	1	0.00
1	400	70	U	0	0.00	1	0.00
2	400	70	U	0	0.00	1	0.00
3	400	70	U	1	0.00	2	0.00
4	400	70	U	1	0.00	3	0.01
5	400	70	U	2	0.00	5	0.01
6	400	70	U	2	0.00	7	0.02
7	400	70	U	6	0.01	13	0.03
8	400	70	U	10	0.02	23	0.05
9	400	70	U	16	0.03	39	0.08
10	400	70	U	31	0.07	70	0.15
11	400	70	U	58	0.12	128	0.27
12	400	70	U	109	0.23	237	0.51
13	475	77	U	154	0.33	391	0.84
14	517	80	U	223	0.48	614	1.32
15	544	79	U	314	0.67	928	1.99
16	565	75	U	402	0.86	1330	2.85
17	583	69	U	494	1.06	1824	3.91
18	597	62	U	542	1.16	2366	5.08
19	610	55	U	622	1.33	2988	6.41
20	621	49	U	776	1.67	3764	8.08
21	632	44	U	850	1.82	4614	9.90
22	641	39	U	952	2.04	5566	11.94
23	650	35	L	1038	2.23	6604	14.17
24	659	33	L	1109	2.38	7713	16.55
25	667	30	L	1140	2.45	8853	19.00
26	674	29	L	1156	2.48	10009	21.48
27	681	27	L	1179	2.53	11188	24.01
28	688	26	L	1266	2.72	12454	26.72
29	695	25	L	1302	2.79	13756	29.52
30	701	24	S	1326	2.85	15082	32.36
31	708	23	S	1354	2.91	16436	35.27
32	714	22	S	1404	3.01	17840	38.28
33	720	22	S	1431	3.07	19271	41.35
34	725	21	S	1440	3.09	20711	44.44
35	731	21	S	1510	3.24	22221	47.68
36	737	20	S	1497	3.21	23718	50.89
37	742	20	S	1535	3.29	25253	54.19
38	748	20	S	1533	3.29	26786	57.48
39	753	20	S	1472	3.16	28258	60.63
40	759	19	S	1456	3.12	29714	63.76
41	764	19	S	1514	3.25	31228	67.01
42	770	19	S	1493	3.20	32721	70.21
43	776	20	S	1386	2.97	34107	73.18
44	782	20	S	1394	2.99	35501	76.18
45	788	20	A	1376	2.95	36877	79.13
46	794	20	A	1266	2.72	38143	81.84
47	800	21	A	1254	2.69	39397	84.54
48	807	21	A	1220	2.62	40617	87.15
49	814	22	A	1004	2.15	41621	89.31

Table Continues

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
50	822	23	A	1008	2.16	42629	91.47
51	830	25	A	970	2.08	43599	93.55
52	839	26	A	806	1.73	44405	95.28
53	849	29	A	656	1.41	45061	96.69
54	860	32	A	502	1.08	45563	97.77
55	874	35	A	397	0.85	45960	98.62
56	890	39	A	335	0.72	46295	99.34
57	911	41	A	162	0.35	46457	99.68
58	942	39	A	98	0.21	46555	99.89
59	990	27	A	42	0.09	46597	99.98
60	990	27	A	7	0.02	46604	100.00

Table B.16
Raw-to-Scaled Score Table and Frequency Distribution: Grade 7 Geography

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	65	U	0	0.00	0	0.00
1	400	65	U	0	0.00	0	0.00
2	400	65	U	1	0.00	1	0.00
3	400	65	U	3	0.01	4	0.01
4	400	65	U	2	0.00	6	0.01
5	400	65	U	6	0.01	12	0.03
6	400	65	U	15	0.03	27	0.06
7	400	65	U	38	0.09	65	0.15
8	400	65	U	66	0.15	131	0.29
9	457	75	U	118	0.26	249	0.56
10	511	80	U	196	0.44	445	0.99
11	546	80	U	257	0.58	702	1.57
12	573	77	U	373	0.84	1075	2.41
13	595	70	L	463	1.04	1538	3.44
14	613	63	L	547	1.23	2085	4.67
15	629	56	L	641	1.44	2726	6.11
16	643	50	L	709	1.59	3435	7.69
17	656	44	L	803	1.80	4238	9.49
18	667	40	L	831	1.86	5069	11.35
19	678	37	L	955	2.14	6024	13.49
20	689	35	L	1071	2.40	7095	15.89
21	699	33	L	1143	2.56	8238	18.45
22	709	32	S	1317	2.95	9555	21.41
23	718	31	S	1311	2.94	10866	24.34
24	728	30	S	1581	3.54	12447	27.89
25	737	30	S	1598	3.58	14045	31.47
26	746	29	S	1761	3.95	15806	35.41
27	755	29	S	1797	4.03	17603	39.44
28	764	29	S	2000	4.48	19603	43.92
29	773	28	S	2032	4.55	21635	48.47
30	782	29	S	2078	4.66	23713	53.13
31	792	29	S	2176	4.88	25889	58.00
32	801	29	S	2227	4.99	28116	62.99
33	811	30	S	2139	4.79	30255	67.78
34	822	31	S	2204	4.94	32459	72.72
35	833	32	S	2060	4.62	34519	77.34
36	844	33	S	2085	4.67	36604	82.01
37	857	35	A	1807	4.05	38411	86.06
38	871	37	A	1751	3.92	40162	89.98
39	886	39	A	1379	3.09	41541	93.07
40	904	41	A	1134	2.54	42675	95.61
41	926	41	A	874	1.96	43549	97.57
42	953	37	A	548	1.23	44097	98.80
43	990	27	A	349	0.78	44446	99.58
44	990	12	A	156	0.35	44602	99.93
45	990	12	A	32	0.07	44634	100.00

Table B.17
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 U.S. History

Raw Score	Scaled Score	CSEM	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	400	76	U	0	0.00	0	0.00
1	400	76	U	1	0.00	1	0.00
2	400	76	U	0	0.00	1	0.00
3	400	76	U	3	0.01	4	0.01
4	400	76	U	6	0.01	10	0.02
5	400	76	U	14	0.03	24	0.05
6	400	76	U	31	0.07	55	0.13
7	400	76	U	44	0.10	99	0.23
8	400	76	U	109	0.25	208	0.47
9	460	76	U	198	0.45	406	0.93
10	519	80	U	295	0.67	701	1.60
11	552	79	U	369	0.84	1070	2.44
12	575	73	U	530	1.21	1600	3.65
13	593	66	U	681	1.55	2281	5.20
14	608	57	U	785	1.79	3066	6.99
15	621	50	U	941	2.14	4007	9.13
16	633	43	L	1003	2.29	5010	11.42
17	644	38	L	1088	2.48	6098	13.90
18	654	34	L	1110	2.53	7208	16.42
19	663	31	L	1196	2.73	8404	19.15
20	672	29	L	1212	2.76	9616	21.91
21	681	27	L	1308	2.98	10924	24.89
22	688	26	L	1325	3.02	12249	27.91
23	696	25	L	1453	3.31	13702	31.22
24	704	24	S	1476	3.36	15178	34.59
25	711	23	S	1510	3.44	16688	38.03
26	718	23	S	1534	3.50	18222	41.52
27	725	23	S	1620	3.69	19842	45.21
28	732	22	S	1651	3.76	21493	48.98
29	739	22	S	1711	3.90	23204	52.87
30	747	22	S	1703	3.88	24907	56.76
31	754	22	S	1699	3.87	26606	60.63
32	761	23	S	1765	4.02	28371	64.65
33	769	23	S	1701	3.88	30072	68.52
34	777	24	S	1726	3.93	31798	72.46
35	785	25	S	1802	4.11	33600	76.56
36	794	26	S	1764	4.02	35364	80.58
37	804	27	S	1662	3.79	37026	84.37
38	815	30	S	1515	3.45	38541	87.82
39	827	33	A	1413	3.22	39954	91.04
40	841	37	A	1305	2.97	41259	94.02
41	858	41	A	1061	2.42	42320	96.43
42	880	46	A	775	1.77	43095	98.20
43	911	46	A	501	1.14	43596	99.34
44	967	36	A	218	0.50	43814	99.84
45	990	36	A	71	0.16	43885	100.00

Table B.18
Composite Score Frequency Distribution: Grade 5 Writing

Composite Score	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
15	U	647	0.36	647	0.36
16	U	12	0.03	659	0.39
17	U	68	0.15	727	0.54
18	U	27	0.06	754	0.60
19	U	12	0.03	766	0.63
20	U	111	0.25	877	0.88
21	U	26	0.06	903	0.93
22	U	74	0.16	977	1.10
23	U	81	0.18	1058	1.28
24	U	74	0.16	1132	1.44
25	U	153	0.34	1285	1.79
26	L	80	0.18	1365	1.96
27	L	173	0.39	1538	2.35
28	L	59	0.13	1597	2.48
29	L	295	0.66	1892	3.14
30	L	219	0.49	2111	3.63
31	L	2553	5.69	4664	9.31
32	L	764	1.70	5428	11.02
33	L	148	0.33	5576	11.35
34	L	1552	3.46	7128	14.80
35	L	960	2.14	8088	16.94
36	S	1736	3.87	9824	20.81
37	S	1183	2.64	11007	23.45
38	S	1949	4.34	12956	27.79
39	S	974	2.17	13930	29.96
40	S	1012	2.25	14942	32.22
41	S	2087	4.65	17029	36.87
42	S	1224	2.73	18253	39.60
43	S	2911	6.49	21164	46.08
44	S	1685	3.76	22849	49.84
45	S	2986	6.65	25835	56.49
46	S	1562	3.48	27397	59.97
47	S	380	0.84	27777	60.82
48	S	6180	13.77	33957	74.59
49	S	1452	3.23	35409	77.82
50	S	2260	5.04	37669	82.86
51	S	949	2.11	38618	84.97
52	S	1533	3.42	40151	88.39
53	S	368	0.82	40519	89.21
54	A	392	0.87	40911	90.08
55	A	847	1.89	41758	91.97
56	A	762	1.70	42520	93.67
57	A	631	1.41	43151	95.07
58	A	272	0.61	43423	95.68
59	A	398	0.89	43821	96.57
60	A	1541	3.43	45362	100.00

Table B.19
Raw-to-Scaled Score Table and Frequency Distribution: Grade 8 Writing

Composite Score	Level	Frequency	Percent	Cumulative Frequency	Cumulative Percent
15	U	421	0.00	421	0.00
16	U	104	0.25	525	0.25
17	U	11	0.03	536	0.27
18	U	43	0.10	579	0.38
19	U	0	0.00	579	0.38
20	U	18	0.04	597	0.42
21	U	50	0.12	647	0.54
22	U	38	0.09	685	0.63
23	U	41	0.10	726	0.73
24	U	45	0.11	771	0.83
25	L	47	0.11	818	0.95
26	L	36	0.09	854	1.03
27	L	114	0.27	968	1.30
28	L	92	0.22	1060	1.52
29	L	85	0.20	1145	1.72
30	L	175	0.42	1320	2.14
31	L	146	0.35	1466	2.49
32	L	1790	4.26	3256	6.75
33	L	18	0.04	3274	6.79
34	L	768	1.83	4042	8.62
35	L	801	1.90	4843	10.52
36	S	740	1.76	5583	12.29
37	S	827	1.97	6410	14.26
38	S	551	1.31	6961	15.57
39	S	1076	2.56	8037	18.13
40	S	605	1.44	8642	19.57
41	S	1409	3.35	10051	22.92
42	S	1077	2.56	11128	25.49
43	S	1077	2.56	12205	28.05
44	S	2055	4.89	14260	32.94
45	S	1446	3.44	15706	36.38
46	S	2416	5.75	18122	42.14
47	S	639	1.52	18761	43.66
48	S	1900	4.52	20661	48.18
49	S	12630	30.06	33291	78.23
50	S	1210	2.88	34501	81.11
51	S	1653	3.94	36154	85.05
52	S	619	1.47	36773	86.52
53	S	492	1.17	37265	87.69
54	A	479	1.14	37744	88.83
55	A	561	1.34	38305	90.17
56	A	568	1.35	38873	91.52
57	A	983	2.34	39856	93.86
58	A	338	0.80	40194	94.66
59	A	354	0.84	40548	95.51
60	A	1888	4.49	42436	100.00